VOLUME 23 MARCH, 1959 NUMBER 1

The Relationship Between Level of ESP Scoring and Student Class Grade ..... *Margaret Anderson* 1

Three-Tier Placement PK ..... *W. E. Cox* 19

Some Avoidable Misconceptions in Parapsychology *J. B. Rhine* 30

Three New ESP Test Machines and Some Preliminary Results ..... *W. C. Stewart* 44

Terminal Salience With Multiple Digit Targets *Carroll B. Nash and Martin G. Durkin* 49

Some Comments on the Controversy Regarding Dowsing Rods ..... Carl von Klinckowstroem 54

Comments and Letters

Statistical Evaluations Used in the Schmeidler and McConnell Book (*T. N. E. Greville*) (*R. H. Thouless*) 64

Book Reviews

The Integration of Human Knowledge by Oliver L. Reiser *Henry Margenau* 66

The World of Psychic Phenomena by F. S. Edsall *Sara R. Feather* 68

Psychical Research by Reginald Omez ..... *W. S. Cox* 68

Parapsychological Abstracts 71

Announcements by the Parapsychological Association 76

Call for Papers for Annual Convention

Glossary 77

IENTIFIC QUARTERLY DEALING WITH EXTRASENSORY PERCEPTION, THE PSYCHOKINETIC EFFECT, AND RELATED TOPICS

DUKE UNIVERSITY PRESS
DURHAM, N. C.

The Journal *of* Parapsychology

EDITORIAL STAFF

J. G. PRATT, L. E. RHINE, and W. G. ROLL, *Editors*
DOROTHY H. POPE, *Managing Editor*
J. A. GREENWOOD and T. N. E. GREVILLE, *Statistical Editors*

This JOURNAL, published on the thirtieth day of March, June, September, and December, is devoted primarily to the original publication of experimental results and other research findings. In addition, articles presenting reviews of the literature relevant to parapsychology, criticisms of published work, proposals for research, theoretical and philosophical discussions, reports of spontaneous cases, and new methods of mathematical analysis will be published as JOURNAL space allows. The editors welcome comments, inquiries, and suggestions about the JOURNAL and any of its contents.

Contributions submitted for publication and all editorial communications should be addressed to the Managing Editor. Correspondence with the editors is advised before submitting articles other than reports of experimentation. Since it does not forward manuscripts by registered mail, the JOURNAL cannot guarantee that they will not be lost in transit. Contributors are therefore urged to keep copies of their papers. All manuscripts should be typewritten, double spaced. References should be given in the form adopted by the JOURNAL. Contributions will be acknowledged immediately upon receipt, and the editors will communicate further regarding each manuscript as soon as possible, generally within two months. Articles will be published in the order of final acceptance insofar as content balance in the JOURNAL will allow.

Reprints may be ordered when the proof is returned. The JOURNAL will bear one-half the cost of reprints up to two hundred copies.

The subscription price is $6.00 a year; single current numbers can be obtained for $1.50 each. The rate for back volumes is $7.00; for single back numbers, $1.75. All remittances should be made payable to the JOURNAL OF PARAPSYCHOLOGY. Missing numbers will be supplied free when lost in the mails if written notice is given within one month of the date of issue. Correspondence concerning subscriptions, change of address, back numbers, and other business communications should be addressed to the Managing Editor, Dorothy H. Pope, College Station, Durham, N. C.

Subscriptions may be obtained in Great Britain through the Cambridge University Press, Bentley House, Euston Road, London, N.W.1, England.

Entered as second class matter May 21, 1942, at the post office at Durham, N. C., under the act of March 3, 1879.

DUKE UNIVERSITY PRESS

The Journal *of*
Parapsychology

VOLUME 23 MARCH, 1959 NUMBER 1

THE RELATIONSHIP BETWEEN LEVEL OF ESP SCORING AND STUDENT CLASS GRADE

By MARGARET ANDERSON[1]

ABSTRACT: In this report the author examined the relationship between students' ESP scoring and their school grades. As material for the analyses, she used the data from ten research studies carried out with students at nine different grade levels. A statistically significant relationship was found between the ESP scores and the grades: the students who received high grades, A and B, tended to average above mean chance expectancy in their ESP tests, while students with low grades, D and E, tended to have scores below chance. The C students did not furnish any evidence of ESP.

When grades and attitudes were compared, it was found that there was a definite relationship between the teachers' attitudes toward the students and the grades they gave them; and there was a less marked association between the students' attitudes toward the teachers and the grades they received.

The most striking negative ESP deviation was contributed by a group of students with D and E grades who disliked their teachers; whereas students with D and E grades who registered a favorable attitude scored close to mean chance expectation. This finding raises the question whether these students did not unconsciously react against learning as they reacted against the ESP targets, rather than being simply less capable than the students with better grades. The possible implications of this research for educational psychology are discussed.—Ed.

INTRODUCTION

ONE WAY to bring ESP to a point of understanding is to trace out its correlates in other areas related to it. Since 1956, assisted by my colleague, Miss Rhea White of the Duke Parapsychology Laboratory, I have conducted a series of 13 experimental studies

[1] I wish to express my gratitude to the 39 teachers and the 1,228 subjects who made this study possible as well as to Dr. J. B. Rhine, Miss Rhea White, and Mrs. Betsy Moser who gave me valuable assistance.

I was supported in this work in part by a Ralph Drake Perry Fellowship.

on the relationship between teacher-pupil attitudes and level of ESP scoring (**1**, **2**, **3**, **4**, **5**, **6**). In 10 of these studies, I collected data pertinent to the primary question in this report, the relationship between ESP scoring and student academic achievement. The measure of this achievement has been the grade—A, B, C, D, or E—which the student was receiving at the time of the ESP test in the class in which this test was conducted.

The pupil's academic achievement might be expected to be closely related to the teacher's and pupil's attitudes toward each other. One might suppose that grades and attitudes might overlap, particularly those of the teacher for the pupil. Another objective of the research, therefore, was to ascertain if and to what extent positive attitudes of each for the other were related to high grades, and negative attitudes to low grades, and each to the level of ESP scoring.

Procedure

All of the studies have been tests for clairvoyance with the exception of two in which tests for precognition were given. The original school experiment (**2**), now referred to as A-W 1, was conducted in seven academic classes at the secondary school level in seven different states. The experiment was then successfully repeated in two fifth- and two sixth-grade classes in Wheaton, Illinois (**1**). This, in turn, was followed by a confirmatory series of the A-W 1 investigation (**3**). None of the 10 studies (seven of which are included in this report) which have been carried out since the first three A-W experiments were intended as repetitions of the A-W series. They have been of an exploratory nature, a searching in various directions for departures and variants of possible value.

Among these 10 researches, the student's academic grades were obtained in seven which therefore supply the data for this report together with those from the first three A-W series where the grades were also recorded.

The 10 studies reported here are as follows: the first three A-W studies; four high-school academic classes in a boys' training or reform school in Iowa; four high-school academic classes in Hillsboro, North Carolina; four sophomore English classes taught by Mr. Claude Burns in San Diego, California; one senior high-school

and two junior-college mathematics classes of Miss Sally Springer in Moberly, Missouri; six non-academic high-school classes in four different states; one sixth and two fifth grades in Wheaton, Illinois; two fourth-grade classes in Wheaton, Illinois, tested by the teacher in collaboration with the principal, Mrs. Mary Davies.

The methods of procedure for these 10 studies have previously been published in detail, (3, 6). I am therefore dispensing with any description of procedure other than this: each teacher was asked to answer the question: If you could form your ideal group for this class would you include this student?—and to indicate the grade the student was receiving at the time of the test.

Results

Class Grades and ESP Scores

The ESP experiments involved 39 teachers of 44 different classes in 9 different grade levels. There were 1,228 subjects, who completed 6,388 runs or 159,700 trials. Of the total number of subjects tested, the 263 receiving A grades gave an average run score of 5.16 in their ESP tests, with 5.00 the mean chance expectation (M.C.E.). Three hundred and sixty-nine B subjects averaged 5.07; 408 C subjects, 4.98: 145 D subjects, 4.84: and the 43 E, or failing subjects, 4.77 (see Table 1).

The 632 A and B students combined gave a total of 16,666 hits on 3,264 runs of ESP tests. This gives a positive deviation of 346 which has a significant critical ratio (CR) of 3.03. The 188 D and E students combined gave a total of 4,741 hits on 983 ESP test runs. This gives a negative deviation of 174 which also has a significant CR of 2.84. The CR of the difference between the two levels of ESP scoring is 4.02 ($P = .00006$).

In order to determine the consistency of the individual scoring above and below M.C.E. with the combined high and the combined low grades, the results were arranged in a two-by-two contingency table. The chi-square is 8.52, which, with one degree of freedom, has a significant P-value of .003 (see Table 2).

Teacher Attitude, Class Grades, and ESP Scores

The relationship between teacher attitude, student grade, and ESP score is shown in Table 3 and Figure 1.

Table 1*

ESP Results as Related to Student Grade

Study	No. Classes	No. Teachers	No. Subjects	No. Runs	Grade A		Grade B		Grade C		Grade D		Grade E	
					No. Runs	Av. Run Score	No. Runs	Av. Run Score	No. Runs	Av. Run Score	No. Runs	Av. Run Score	No. Runs	Av. Run Score
A-W 1	7	7	228	1140	175	5.24	320	5.23	360	4.97	230	4.94	55	4.75
A-W 2	7	7	205	1025	225	5.33	255	5.29	350	4.97	145	4.81	50	4.78
Wheaton	4	4	113	446	144	4.92	128	5.17	124	5.00	32	4.63	18	4.78
Total	18	18	546	2611	544	5.19	703	5.24	834	4.97	407	4.87	123	4.78
Training School	4	4	69	345	50	5.26	115	5.05	155	5.00	25	4.80		
Hillsboro	4	4	105	525	135	5.13	135	5.04	165	5.00	40	4.70	50	4.92
San Diego	4	1	93	930	150	5.21	320	4.87	340	5.05	120	4.63		
Missouri	3	1	87	435	70	5.14	125	5.16	150	5.05	75	4.84	15	4.40
Non-Academic	6	6	161	805	165	4.99	320	5.04	240	4.83	65	5.23	15	4.60
Wheaton	3	3	86	348	88	5.10	132	4.84	120	4.96	4	4.75	4	5.00
Wheaton	2	2	81	389	125	5.26	87	4.75	137	4.97	40	4.58		
Total	26	21	682	3777	783	5.14	1234	4.97	1307	4.98	369	4.80	84	4.77
Total for 10 Studies	44	39	1228	6388	1327	5.16	1937	5.07	2141	4.98	776	4.84	207	4.77

CR High Grades = 3.03 CR Low Grades = 2.84
$CR_d = 4.02$ ($P = .00006$)

* The data of this study are taken from the experiments reported in "A Survey of Work on ESP and Teacher-Pupil Attitudes." *J. Parapsychol.*, 1958, **22**, 246-69. There are 143 subjects for whom grades were unavailable, and three subjects for whom information was obtained following publication of the above-mentioned report. There are, therefore, 1,228 subjects involved in this study as compared to the 1,368 subjects in the previous report.

The teachers were favorable, or positive, in attitude toward 816 (67%) of the 1,228 subjects in this study as compared to 412 (33%) subjects for whom they indicated an unfavorable, or negative, attitude. Of the 816 liked subjects, 253 were A students, who gave an average ESP score of 5.18. There were 301 students with B grades who gave an average of 5.07; 233 students with grade C, who averaged 5.00; 26 students with D's, who averaged 5.01; and only three students with the failing grade of E, and they averaged 5.60.

Table 2

ESP Scores in Relation to High and Low Grades

	Students with Grades A-B	Students with Grades D-E	Total
Students with ESP Scores Above M.C.E.	300	66	366
Students with ESP Scores Below M.C.E.	273	101	374
	573	167	740

$x^2 = 8.52$ (1 d.f.)
$P = .003$

Table 3

Grade Distribution and ESP Results as Related to Attitude of Teacher

Teacher Attitude	A-Grade Students				B-Grade Students				C-Grade Students			
	N	% of Total Subjects	% of A-grade Subjects	Av. Run Score	N	% of Total Subjects	% of B-grade Subjects	Av. Run Score	N	% of Total Subjects	% of C-grade Subjects	Av. Run Score
Positive	253	21	96	5.18	301	25	82	5.07	233	19	57	5.00
Negative	10	<.5	4	4.76	68	6	18	5.02	175	14	43	4.95

Teacher Attitude	D-Grade Students				E-Grade Students			
	N	% of Total Subjects	% of D-grade Subjects	Av. Run Score	N	% of Total Subjects	% of E-grade Subjects	Av. Run Score
Positive	26	2	17	5.01	3	<.5	7	5.60
Negative	119	10	83	4.80	40	3	93	4.98

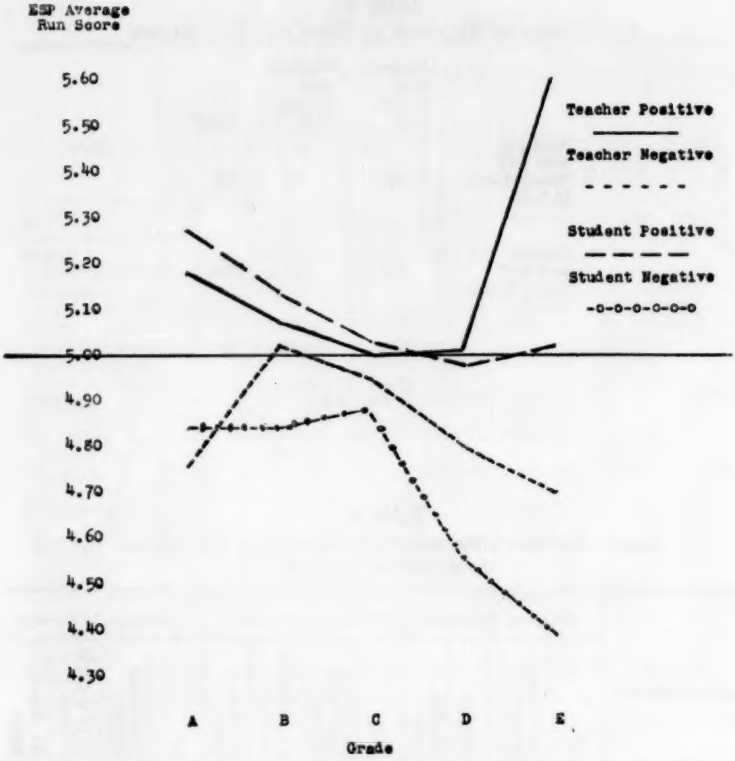

FIG. 1. Teacher's and student's attitude as related to student's grade and level of ESP scoring.

Of the 412 subjects for whom the teachers indicated a negative attitude, 10 were A students averaging 4.76; 68 were B students, who averaged 5.02; 175 were C students, averaging 4.95; 119 were D students, who averaged 4.80; and 40 were E students who averaged 4.98.

The percentages of pupils receiving given grades according to teacher attitude are presented in Table 3. Of the total 1,228 grades where the teacher was positive in attitude toward the student, 21% were A, 25% were B, 19% were C, 2% were D, and less than .5% E. By contrast, where the teacher was negative in attitude, the

grade distribution was less than .5% A, 6% B, 14% C, 10% D, and 3% E.

From the standpoint of the distribution of the particular grade only (as is also shown in Table 3), the results reveal that, of the 263 students getting A grades, 96% were given positive ratings by the teachers as contrasted to 4% who were given negative ratings. The teachers gave positive ratings to 82% of the 369 B students as contrasted to the negative ratings they gave to 18%. The 408 C students were divided into 57% who received positive ratings and 43% who were given negative ratings. Of the 145 D students, 17% were rated positively as against 83% who had negative ratings; and the 43 E students were divided into 7% with positive ratings and 93% with negative ratings. For a graphic representation of these percentages the reader may refer to Figures 2 and 3.

In order to determine the consistency of individual scoring above and below M.C.E. with the combined high and the combined low grades and teacher attitude, the results were arranged in a two-

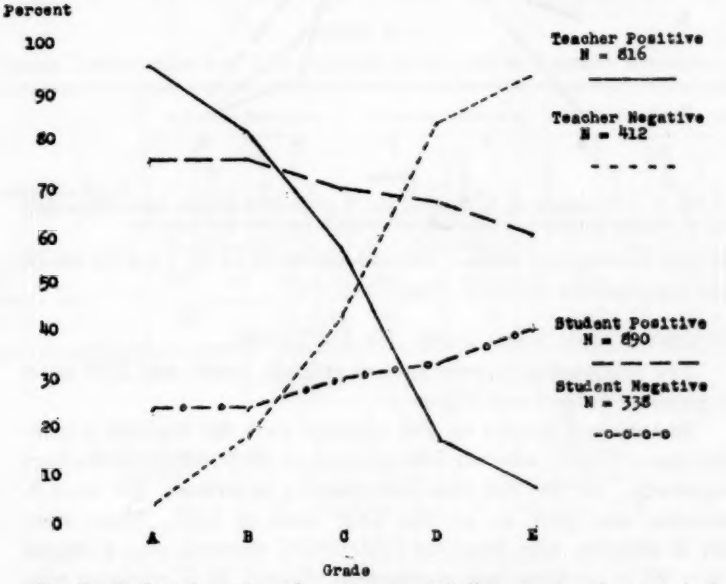

FIG. 2. Students in each grade group separated (by percentages) on the basis of student and teacher attitude.

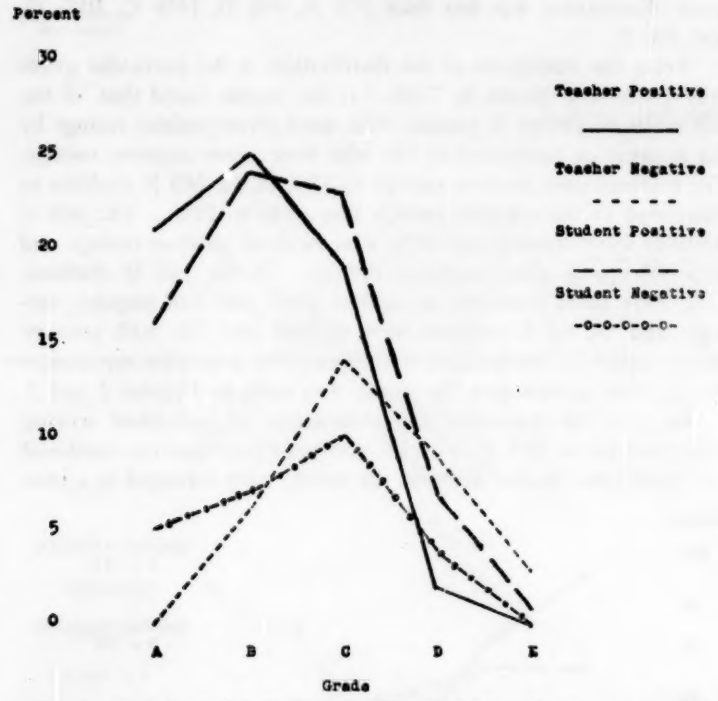

FIG. 3. Percentages of A, B, C, D, and E in the total grades, separated according to teacher positive and negative and student positive and negative.

by-two contingency table. The chi-square is 11.55 (1 d.f.) which has a probability of .0007 (see Table 4).

Student Attitude, Class Grade, and ESP Score

The relationship between student attitude, grade, and ESP score is given in Table 5 and Figure 1.

Seventy-two percent or 890 students gave the teachers a positive rating (10-7) whereas 338 students, or 28%, rated the teachers negatively. Of the 890 who were positive in attitude, 201 were A students, who gave an average ESP score of 5.27. There were 281 B students, who averaged 5.13; 285 C students, who averaged 5.03; 97 D students, who averaged 4.98; and 26 E students, who averaged 5.02.

Table 4

STUDENT ESP SCORE LEVEL IN RELATION TO HIGH-LOW GRADES AND TEACHER'S ATTITUDE TOWARD STUDENT

	Teacher: Pos. Grades A-B	Teacher: Neg. Grades D-E	Total
Students with ESP Scores Above M.C.E.	270	53	323
Students with ESP Scores Below M.C.E.	232	88	320
	502	141	643

$x^2 = 11.55$ (1 d.f.)
$P = .0007$

Table 5

GRADE DISTRIBUTION AND ESP RESULTS AS RELATED TO STUDENT ATTITUDE

Student Rating of Teacher	A-GRADE STUDENTS				B-GRADE STUDENTS				C-GRADE STUDENTS			
	N	% of Total Subjects	% of A-grade Subjects	Av. Run Score	N	% of Total Subjects	% of B-grade Subjects	Av. Run Score	N	% of Total Subjects	% of C-grade Subjects	Av. Run Score
10-7	201	16	76	5.27	281	23	76	5.13	285	23	70	5.03
6-0	62	5	24	4.84	88	8	24	4.84	123	10	30	4.88

Student Rating of Teacher	D-GRADE STUDENTS				E-GRADE STUDENTS			
	N	% of Total Subjects	% of D-grade Subjects	Av. Run Score	N	% of Total Subjects	% of E-grade Subjects	Av. Run Score
10-7	97	8	67	4.98	26	2	60	5.02
6-0	48	4	33	4.55	17	1	40	4.39

Of the 338 students who gave the teacher a negative rating (6-0), 62 were A students, who had an average ESP score of 4.84. The 88 B students averaged 4.84; the 123 C students averaged 4.88; the 48 D students averaged 4.55; and the 17 E students averaged 4.39.

The distribution of grades from the viewpoint of the student attitude is given in Table 5. When the student was positive in attitude toward the teacher, of the total 1,228 grades, 16% were A, 23% were B, 23% were C, 8% were D, and 2% were E. Conversely, where the student was negative in attitude, the grade distribution was: A, 5%; B, 8%; C, 10%; D, 4% and E, 1%.

An analysis of the distribution of the particular grade only, according to the student attitude, shows that of the 263 students receiving A grades, 76% liked the teacher as contrasted to 24% who did not. Of the 369 B grade students, 76% gave the teacher positive attitude ratings as contrasted to 24% who gave negative ratings. Of the 408 C students, 70% were positive in attitude and 30% were negative; of the 145 D students, 67% were positive and 33% were negative; and of the 43 E students, 60% were positive and 40% were negative. (For graphic representation of these data see Figures 2 and 3.)

In order to determine the consistency of individual scoring above and below M.C.E. divided according to the combined high and the combined low grades and the positive and negative student attitudes, the results were arranged in a two-by-two contingency table. The chi-square is 17.50 (1 d.f.) which has a probability of .00003 (see Table 6).

Teacher Attitude, Student Attitude, Class Grade, and ESP Scores

Figure 4 graphically indicates that where teacher and student were positive in attitude toward each other, the level of ESP scoring was above chance regardless of the grade the student was receiving. (I would remind the reader that the ESP average of 5.60 for the E grade was contributed by only three subjects.) When teacher and student shared negative attitudes toward each other, the level of ESP scoring was below chance regardless of the grade. (Here too I would point out that the ESP average of 4.27 for the A grade was contributed by only two subjects.)

Table 6

STUDENT ESP SCORE LEVEL IN RELATION TO HIGH-LOW GRADES AND STUDENT'S ATTITUDE TOWARD TEACHER

	Student Grades: Pos. A-B	Neg. D-E	Total
Students with ESP Scores Above M.C.E.	241	15	256
Students with ESP Scores Below M.C.E.	202	44	246
	443	59	502

$\chi^2 = 17.50$ (1 d.f.)
$P = .00003$

A closer examination of the results of the ESP scoring and the distribution of grades according to the attitude of the teacher and the student toward each other is presented in Table 7. The teacher registered a positive attitude toward 636 subjects, who in turn gave the teacher a positive rating. Of those 636, 435 were A or B students who, on 2,223 runs, gave an ESP score of 11,561 hits. This gives a positive deviation of 446 which has a corresponding CR of 4.62 with a P-value of .000,004.

Of the 158 students who registered a negative attitude for the teacher and who were given a negative rating by the teacher, 61 were D or E students. On 325 runs they gave an ESP score of 1,471 hits which represents a negative deviation of 154, the CR of which is 4.27 with a P of .00002.

In order to determine the consistency of individual scoring above and below M.C.E. divided according to the combined high grades and mutually positive attitudes and combined low grades and mutually negative attitudes, the results were arranged in a two-by-two contingency table. The chi-square is 15.82 (1 d.f.) which gives a P-value of .00006 (see Table 8).

I was fortunate in being able to secure data from an independent study with which to compare these findings. Mr. van Busschbach carried out his fifth- and sixth-grade GESP tests in America in the

Table 7

ESP Results and Grade Distribution as Related to Attitude of Teacher and Student Toward Each Other

Student Rating of Teacher	Grade	Teacher Positive Toward Student					Teacher Negative Toward Student				
		N	Percent of the Grade	No. Runs	Av. Run Score	CR	N	Percent of the Grade	No. Runs	Av. Run Score	CR
10-9	A or B	278	44	1393	5.18	3.29	25	4	122	5.02	
	C	96	23	505	5.07		54	13	273	4.90	
	D or E	14	7	70	5.16		52	28	278	4.90	
8-7	A or B	157	25	830	5.23	3.31	22	3	112	5.23	
	C	80	20	396	4.99		55	14	274	5.14	
	D or E	11	6	53	5.11		46	24	237	4.99	
10-7	A or B	435	69	2223	5.19	4.62	47	7	234	5.12	
	C	176	43	901	5.03		109	27	547	5.02	
	D or E	25	13	123	5.18		98	52	515	4.94	
6-0	A or B	119	19	648	4.85		31	5	159	4.79	
	C	57	14	330	4.91		66	16	363	4.85	
	D or E	4	2	20	4.30		61	33	325	4.52	4.27

late spring of 1956 (14). Miss White and I had the opportunity of assisting him in some of his tests, primarily in the town of Burlington, North Carolina. Through acquaintance with the school officials, and with assistance from Miss White, I was able to go through the files when school opened in the fall of 1956 and record the preceding year's final class grade for each of the students still in the school who had participated in the van Busschbach experiment. Since this experiment consisted of 30 trials per subject, the results were presented in percentages, with 20% being M.C.E. The findings showed that of 341 subjects, 68 who were A students averaged 21.52%; 153 B students averaged 20.89%; 98 C students averaged 20.58%; and 22 D students averaged 18.78%. There were no failing (E) subjects.

Table 8

STUDENT ESP SCORE LEVEL IN RELATION TO HIGH-LOW GRADES WHEN STUDENT AND TEACHER HAD SAME ATTITUDE TOWARD EACH OTHER

	Teacher: Pos. / Student: 10-7 / Grades: A-B	Teacher: Neg. / Student: 6-0 / Grades: D-E	Total
Students with ESP Scores Above M.C.E.	219	14	233
Students with ESP Scores Below M.C.E.	181	40	221
	400	54	454

$\chi^2 = 15.82$ (1 d.f.)
$P = .00006$

DISCUSSION

The results of this inquiry have shown a significantly consistent relation between level of ESP scoring and teacher's estimation of student academic achievement. The question which then follows is: Are these results more than merely a reflection of the previously reported findings of a correlation between ESP scoring and teacher-pupil attitudes?

At the outset it was pointed out that while an overlapping element between grades and attitudes was anticipated, particularly as registered by the teacher, they could not be considered as identical. It would be perhaps more accurate to suggest that grades and teachers' attitudes are both measures of much the same reality. In this particular study, it will be recalled that the teacher questionnaire gives a two-point rating while the class grade gives a five-point evaluation.

The degree of overlap may be gauged roughly by the fact that 68% of the students who were liked by the teacher received high grades (A or B) as compared to 3% who received low grades (D or E), the C grades not being under consideration here. Only 19% of the students who were given a negative rating by the

teacher received A or B grades as compared to 39% who received D and E grades.

Student attitude, on the other hand, is by no means so closely consistent with grades. While 54% of the students who registered a positive attitude for the teacher received high grades, as compared to 14% who received low grades, 45% of the students who indicated a dislike for the teacher received high grades and 19% received low grades. As is evident, the grade percentages are not greatly different whether the student likes or dislikes the teacher.

There are obviously many other factors involved in the grade and score correlation. One of the first variables which comes to mind is the intelligence of the pupil. Some experiments have been reported in parapsychological literature which suggest a relationship between I.Q. score and ESP score (**7, 8, 9, 10, 11, 15**). However, the nature of the relationship is clouded by ambiguity since no two investigators have used the same intelligence test and since the studies reported have involved relatively small samples. This variable of pupil intelligence and its relation to ESP scoring is a subject for further study.

Other variables would emerge from differences in school policies, teacher personalities, and standards and uses of grades. For example, in some schools the lowering of a grade is used as a disciplinary measure. Whether the child's behavior or the teacher's personality initiates and determines the nature of the rapport in the classroom relationship—that is, which is cause and which effect—is difficult to ascertain, but the importance of attitude cannot be minimized.

An interesting aspect of the scoring concerns the C student, often referred to as the academically "average" student. The 408 C students gave an average run score of 4.98, just about chance. In an effort to determine if there were extrachance results in the ESP scores of the C students, the data were arranged in a one-by-two contingency table on the basis of above and below M.C.E. scoring in the ESP tests. The resulting chi-square was insignificant. One can only wonder if these student-subjects in the class environment presented in these studies were not less motivated—that for them static rather than dynamic conditions prevailed in the classroom.

The parallel between high and low class grades and high and low ESP scores is, of course, extrachance. What can we make of this and the fact that psi goes negatively? Do the D's and E's suggest that these are *not just low* grades but perhaps *negative* grades?

Turning to the data on student attitude for more light on this question we find that the correlation of low grades and ESP scores turns out to be clearly divided by this breakdown. When the student is positive in attitude toward the teacher but receives a low grade of D or E, his ESP results are close to mean chance expectation. However, when the student is negative in attitude toward the teacher and receives low grades, there is a large negative deviation in the ESP scores. It will be seen in Figure 1 that it was those subjects who obtained D and E grades and also registered disapproval of the teacher, who contributed the principal negative deviation. This group of 65 students gave a highly significant deviation and, in fact, registered the largest average deviation from mean chance expectation of any comparably large block of students. It is obvious, then, that this group deserves special consideration in the interpretation of the results. Fortunately, the other D and E students, those who registered a liking for the teacher, provide a reasonably good control on the group of 65 under discussion; and, it will be noted, these averaged close to the chance mean in their ESP tests.

Both from the point of view of interest in the factors governing ESP performance and in the consideration of the educational process concerned with the production of grades of D and E, this grouping of students raises interesting and important questions. The ESP results indicate a statistically significant ability to avoid or miss the target. While tentative hypotheses do exist to explain this aspect frequently encountered in ESP tests, I do not feel that there is sufficient evidence available at this time to interpret why or how the subject misses the target. There is little doubt, however, that this avoidance reaction is almost certainly unconscious and thus prompts the question: *To what extent is this analogous to the student's learning performance which resulted in his grades of D and E?* To what degree, if any, is the D or E student of this group unconsciously thwarting or thwarted in his conscious effort to learn

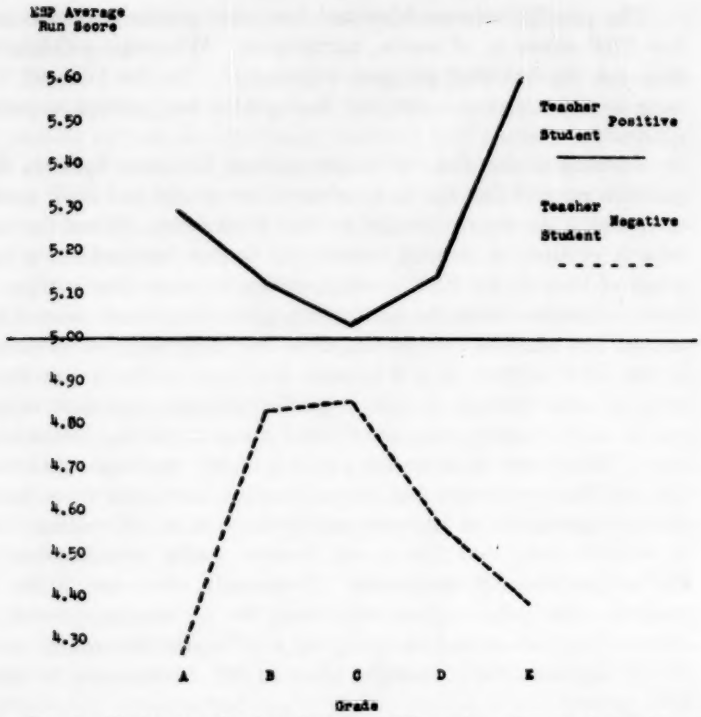

FIG. 4. Grades as related to level of ESP scoring when teacher and pupil have the same attitude toward each other.

and perform well in the class? In the case of the ESP tests there is the advantage that both performance and direction of deviation are registered. Class grades do not tell so much of the story, but a further study might well bring out the facts if it were to be made, and perhaps something of the answer may be found by taking the measure of the student's intelligence for correlation with the ESP scores and class grades. If we had complete data on the I.Q.'s of the groups under discussion, we might more easily understand what is going on in the educational process with these D's and E's who dislike the teacher. There is much to suggest that the D and E students who liked the teacher and who scored at the "chance" level in ESP tests would give a distinctly lower I.Q. average than the D

and E students who disliked the teacher and who gave the remarkably significant negative deviation in ESP tests. This latter group might be expected to show some I.Q.'s comparable to the A and B students. Should further study based on data adequate for this additional analysis bear out such an expectation, the implication both for ESP and perhaps even more for educational psychology could be far-reaching. In any case, the question is clearly raised by the present results as to whether the teaching function may not have to deal with a comparable factor, if not indeed the same factor, the parapsychologist is encountering in his test program under the name of psi-missing.

References

1. Anderson, M. Clairvoyance and teacher-pupil attitudes in fifth and sixth grades. *J. Parapsychol.*, 1957, **21**, 1-11.
2. Anderson, M., and White, R. Teacher-pupil attitudes and clairvoyance test results. *J. Parapsychol.*, 1956, **20**, 141-57.
3. ———. A further investigation of teacher-pupil attitudes and clairvoyance test results. *J. Parapsychol.*, 1957, **21**, 81-97.
4. ———. ESP score level in relation to students' attitude toward teacher-agents acting simultaneously. *J. Parapsychol.*, 1958, **22**, 20-29.
5. ———. The relation between changes in student attitudes and ESP scoring. *J. Parapsychol.*, 1958, **22**, 167-75.
6. ———. A survey of work on ESP and teacher-pupil attitudes. *J. Parapsychol.*, 1958, **22**, 246-68.
7. Bond, E. M. General extrasensory perception with a group of fourth- and fifth-grade retarded children. *J. Parapsychol.*, 1937, **1**, 114-22.
8. Humphrey, B. M. ESP and intelligence. *J. Parapsychol.*, 1945, **9**, 7-16.
9. ———. A further study of ESP and intelligence. *J. Parapsychol.*, 1948, **12**, 213-17.
10. Nicol, J. F., and Humphrey, B. M. The exploration of ESP and human personality. *J. Amer. Soc. psych. Res.*, 1953, **47**, 133-78.
11. Rivers, O. B. An exploratory study of the mental health and intelligence of ESP subjects. *J. Parapsychol.*, 1950, **14**, 267-77.
12. van Busschbach, J. G. An investigation of extrasensory perception in school children. *J. Parapsychol.*, 1953, **17**, 210-14.
13. ———. A further report on an investigation of ESP in school children. *J. Parapsychol.*, 1955, **19**, 73-81.

14. ———. An investigation of ESP between teacher and pupils in American schools. *J. Parapsychol.*, 1956, **20**, 71-80.

15. VASSE, P., AND VASSE, C. ESP tests with French first-grade school children. *J. Parapsychol.*, 1958, **22**, 187-204.

Parapsychology Laboratory
Duke University
Durham, North Carolina

THREE-TIER PLACEMENT PK

By W. E. Cox

ABSTRACT: The following exploratory experiment made use of a new apparatus for measuring the PK placement effect. In the ordinary apparatus, the throwing area is divided into left and right sides, one of which is the target. It is hypothetically possible that the subject in such an experiment may be influencing some cubes positively to go to the target side and others negatively to go to the nontarget side. If this is so, the two simultaneous effects would tend to cancel each other out. The three-tiered apparatus used here was designed to avoid this cancelling effect. Cubes were released into the top tier and if they fell into the target area, they were allowed to go down into the middle tier. Otherwise, they were caught and held back. Those falling to the target side of the middle tier were allowed to continue to a bottom tier divided into 48 squares alternately labelled as target areas.

In this experiment, 24,000 cubes were released into the top tier. Only those which fell through to the bottom gave significant evidence of placement PK.—Ed.

INTRODUCTION

IT HAS BEEN observed that evidence for PK may be either positive or negative ("psi-missing"). For instance, in a placement PK experiment the subject may attempt in alternate releases to influence the falling objects, usually cubes, to fall into the left and right halves of the throwing surface. If more dice come to rest on the designated sides than are expected by chance a positive score results. But it may also happen that the cubes equally conspicuously *miss* the target, in which case there is a negative deviation from chance. Such cubes as are not influenced by either type of PK are likely to be about equally distributed in the target and nontarget halves of the throwing surface.

It is possible that the positive and negative effects may operate simultaneously. In a release of several cubes the subject may produce a positive PK effect on some of them and a negative effect on others. If these two influences are equally strong, there will be a tendency for the scores to cancel out at chance expectancy. A conventional dice-throwing apparatus offers no way of telling whether an insignificant score is due to absence of PK or to such a cancelling of positive and negative scoring. Even in tests where positive scores are obtained—for example, in the usual type of placement

experiment (1, 2, 3) in which a positive result is produced by an excess of cubes on the target side of the throwing surface—it is conceivable that the other side also may have contained cubes which received a PK influence, but a negative rather than a positive one. If this should be the case, a greater sensitivity to measurable PK forces might result by using a machine with several target areas mounted one above the other and connected in such a way as to allow only "hit" cubes that entered the target areas of the upper tiers to fall the entire distance. Any *negatively* influenced dice would be caught and retained in one of the upper tiers. All *uninfluenced* cubes would be distributed onto the "hit" and "miss" sides of the target areas fairly equally by chance.

It should be noted that this type of machine would be advantageous over conventional machines only if a given cube were continously influenced by a positive (or a negative) PK force from the beginning to the end of its movement. If there were a positive influence on a given cube at one moment, say at the top tier, and a negative at the next, when it entered the tier beneath, the scoring in the bottom target area would not be likely to show any improvement over the scoring on the top. If, however, scoring on the bottom tier were relatively higher than on the top tiers this could be an indication that once a cube is influenced by a positive or negative PK effect, it retains this particular influence for the duration of its movement.

This report describes some tests carried out with an apparatus constructed to discover whether higher PK scores could be obtained in the bottom of three tiers, indicating that there might be such a continuous PK influence on the cubes in a PK placement test.

Apparatus

The apparatus used in these experiments is shown in Figure 1. It had three tiers, of which the top measured 12" × 18"; the middle, 16" × 23"; and the bottom, 16" × 21½". All tiers were walled to prevent the cubes from falling out. The three tiers were divided into equal hit and miss areas. The two upper tiers were tipped to allow the cubes which fell in the target ("hit") area to continue their descent. The bottom tier, which rested on the floor, had a slight incline to slow down the cubes so as to prevent them from

piling up at the farther end of the tier. Both halves of the upper tiers had doors at the lower ends, and at both levels the door on the hit-side was open during a release and the other one was closed. Hit-side dice therefore could descend freely to the bottom tier, while those which entered a miss-side were caught and retained. In the top and middle tiers, the two areas were separated by a wire running down the center of the tier at surface level. Both upper tiers were covered with rubber matting. A central partition, which extended part way up the runways, prevented a cube in one area from entering the other during any rebound.

The two sides of the two upper tiers were labelled A or B. The bottom tier was divided into 48 squares, alternately labelled A and B, likewise separated by wires. This arrangement of the target areas was considered more suitable for the bottom tier because it prevented the possibility that the dice when released from the left side, say, of the middle tier would tend to go into the left side of the bottom tier and thus only duplicate the effects of the middle tier. With the regular alternation of the 48 hit and miss squares it did not matter where, precisely, the dice entered the bottom tier. In order to avoid a similar possibility on the middle tier with regard to the dice that entered from either side of the top tier, a duct with several obstructions was inserted between the two upper tiers.

The apparatus had a V-shaped release trough for the dice, with an electric release mechanism attached by a slack wire to a push button. The trough stood approximately six feet above floor level. Commercial dice were used in the tests, and their faces were disregarded. When released, the dice rolled down a sharply inclined runway, covered with rubber, into the top tier. In the experimental series only those which fell into the target half were allowed to descend into the middle tier through the duct and another connecting runway (see Figure 1). Finally, only the dice which landed on the target side of the middle tier continued into the bottom tier. It took approximately three seconds for all unobstructed dice to complete their course through the machine.

The research reported here is of an exploratory nature and the apparatus was not of the best construction. Though it was fairly sturdy and rested on a concrete floor (covered by a rug), it had some weaknesses. The walls of the two upper tiers proved in some in-

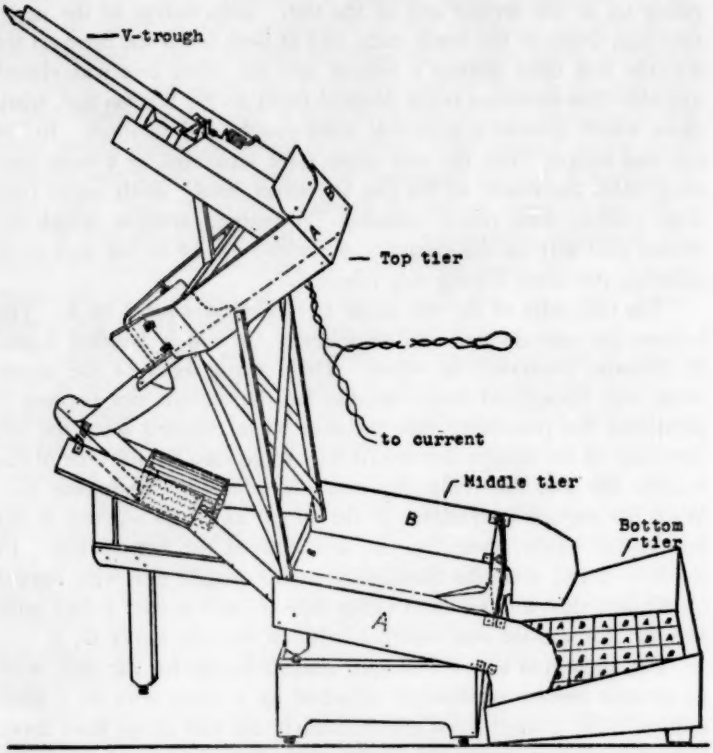

FIG. 1. Drawing of three-tier apparatus. The cubes follow a course from the V-trough on top of the apparatus, down a runway into the top tier (shown here with the trap open on the A-side). Those which are not caught continue down through a corrugated duct and runway into the middle tier. The cubes which pass through the open side of the lower end of this tier fall directly onto the alternately labelled checkerboard surface of the bottom area.

stances not to be high enough, and occasionally a die would bounce out of the apparatus. Sometimes, too, a die halted on the open side of the middle tier, which had a smaller angle of decline than the top tier. When a cube bounced outside from the top tier, it was electrically re-released from the V-trough after all others had been counted and removed from the apparatus. Also at that time any cubes that had stopped on the target side of the middle tier were released once more. Since these were unquestionably "hits" at the

middle level, each was released and re-released manually from the edge of the target side of the top tier until it fell all the way to the bottom tier. Cubes that bounced out from the middle tier were electrically released again from the chute. About half a dozen dice halted or bounded off the upper tiers during a session. None bounced out of the bottom tier. All traps were opened and closed manually.

Experimental Procedure

The research was carried out in the summer of 1958. It comprised two series, each consisting of 500 releases of 24 dice. Five releases constituted a set and there were four sets to each session. Each series consisted of 25 sessions, with a total of 12,000 cube-releases (24,000 in the two experimental series combined). One experimental session took about 35 minutes to complete. One record page was devoted to each session.

A total of 20 subjects took part in the tests. In Series I there were three subjects only, two of whom also participated in Series II. The three subjects in Series I completed 7, 7, and 11 sessions respectively, whereas each of the 19 subjects in Series II did only one or two. The experimenter participated as subject in Series I, but because of his low deviation, he did not continue in this capacity in Series II.

The target areas were regularly alternated throughout the tests. In both series, the target order was determined before the tests began. Each group of five releases (a set) was thrown for the same target side, but the targets for the four sets in each session of Series I were varied in the following manner. Session 1: ABAB; Session 2: BBAA; Session 3: AABB; Session 4: BABA; Session 5: ABBA; Session 6: BAAB. Sessions 7-12, 13-18, and 19-24 repeated this order. Session 25 was like Session 1. In Series II, the four sets in all 25 sessions had the ABAB order.

The subjects were requested to think of the target areas and the associated letter (A or B). Their attention was directed mainly to the bottom tier. This was also the natural focus of attention, since the dice could be watched as they came to rest. The top tier, on the other hand, was out of sight, and the dice fell rapidly through both that and the middle tier. The subjects sat or stood anywhere

they chose, near the machine. One lay down on a couch opposite it. The release button was operated by the subject whenever he wanted to after the trough had been filled. The cubes were placed in the trough by the experimenter, who was the only person in actual contact with the machine during the experiments.

After each release the experimenter recorded the hits on the bottom tier. He calculated the hits on the middle tier (i.e., the dice which passed through it) by adding the hits and misses on the bottom. Similarly, with regard to the top tier, its hits were obtained by adding the hits and misses of the middle tier. The misses on the top tier, which had been caught by the trap doors, were counted as they were picked up, as a check against recording errors.

Before the tests began, the question arose as to whether a cube released from the left side of the top tier might tend (or not tend) toward the same side of the middle tier. In addition to the duct between these two tiers, another precaution against this possibility was taken in Series I. On the top tier the labels bearing the letter A or B were exchanged in a systematic manner decided on in advance, while the left side of the middle tier was always designated as the A side. By this procedure the hit and miss areas of the two upper levels were on the same sides in half, or 250 of the releases, but on opposite sides in the other 250. A comparison of the results produced under these two conditions showed no evidence that having the target and miss areas at the same sides of the top and middle tiers influenced the direction of the cubes, and the latter arrangement therefore was used throughout Series II, with the left sides of both tiers always being A and the right sides B.

Concurrently with Series II, a control series was conducted, using six dice, all of which were made to travel all the way from the V-shaped release trough to the bottom tier by keeping all doors open. It was the purpose of this test to allow a comparison between PK placement with one throwing surface and the new three-tier apparatus. The apparatus thus became the equivalent of a conventional placement device though it provided a longer incline than is customarily employed. The subject released these six dice, in four sets of five releases each, for sides ABAB. The control trials always were conducted immediately prior to the three-tiered main series. Each of the subjects completed 4 sets, totalling 5 runs (i.e., 120

trials) at each session, amounting to 3,000 cube-releases in the entire control series.

Results

The results for the individual series and for both series pooled are given in Table 1. The total released into the top tier in each series equals the 25 sessions of 20 runs each, or 12,000 dice.

Series I yielded no significant CR's. However, the score on the bottom tier was higher than that on the upper tiers. Its deviation was + 49, in the expected direction, with a CR of 1.80. Series II did achieve significance, with a CR of 2.50 (P = .006) at the bottom level. It should be noted that this was accomplished by the uncommonly low total of 3,003 dice. The two series pooled yield a CR of 3.07 for the bottom tier with a P-value of .001. If these results are converted into an average score per run of 24 dice, this becomes 12.48 for the bottom tier, and 12.07 and 11.88 respectively for the middle and top tiers.

The control hit score was within chance expectancy (1,495 for 125 runs; Dev. = − 5).

Since the purpose of the three-tier plan was to provide a means of enhancing the sensitivity of PK measurement, the CR's of the difference between the bottom and top tiers are of particular interest. As listed beneath Table 1, these are 2.12 for Series I; 2.78 for Series II; and 3.24 for the two series pooled.

The data were analyzed for distribution effects in the session

Table 1

PK Placement Results for Each of Three Tiers

Series	Total Cubes Released	Top Tier				Middle Tier				Bottom Tier			
		Hits	Dev.*	Av. Score per 24 Cubes	CR	Hits	Dev.†	Av. Score per 24 Cubes	CR	Hits	Dev.¶	Av. Score per 24 Cubes	CR
I.....	12,000	5,935	− 65	11.87	1.18	2,968	+½	12.00	—	1,533	+ 49	12.39	1.80
II.....	12,000	5,936	− 64	11.88	1.17	3,003	+35	12.14	0.91	1,571	+ 69½	12.56	2.50
Totals...	24,000	11,871	−129	11.88	1.66	5,971	+35½	12.07	0.66	3,104	+118½	12.48	3.07

CRITICAL RATIOS OF THE DIFFERENCE: { Series I, Bottom tier-Top tier = 2.12
Series II, Botton tier-Top tier = 2.78
Series Pooled Bottom tier-Top tier = 3.24

*Equals hits less half of cubes released.
†Equals hits less half of top tier's hits.
¶Equals hits less half of middle tier's hits.

and in the set. On the top and bottom tiers the scores were closer to chance in the second half of the session than in the first, whereas the middle tier showed a slight improvement in the second half. None of the differences were statistically significant.

The distributions of scores within the set are shown in Figure 2. Though they also are insignificant, there is a close resemblance between the curves representing the scores on the bottom and middle tiers. The top tier curve, on the other hand, is almost a mirror image of the others.

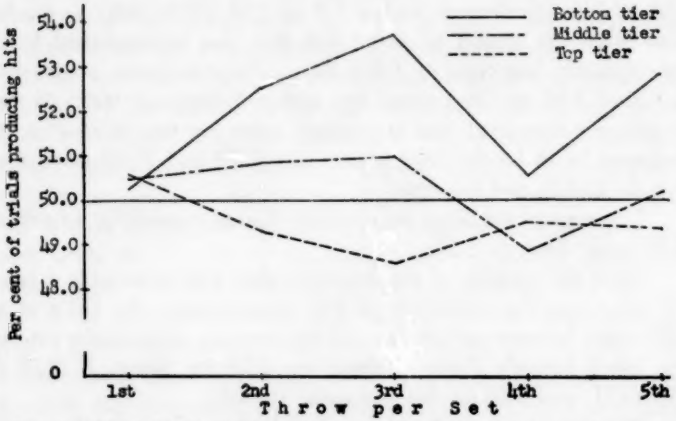

FIG. 2. Graph of hit distributions among the 5 throws per set, for Series I and II (pooled) by tiers.

DISCUSSION

According to the hypothesis outlined in the introduction, the positively influenced cubes should be the more likely to descend completely through to the bottom tier, with the negatively influenced ones being caught at the top or middle tiers. This is, in fact, what seemed to happen, as the deviation in the bottom tier was much higher than at the top or middle. With regard to the scores on the middle tier, according to the hypothesis they should fall between the top and bottom tier deviations, which is in fact what occurred (see Table 1).

The present results thus fit the hypothesis very well. Nevertheless, we must remember that the bottom tier probably also had a

psychological advantage over the first. The scoring on this tier was of greatest interest to the experimenter and he did not try to conceal this attitude from the subjects. Hence an alternative hypothesis is that the subject exercised more PK at the bottom than at the upper levels. In future experiments with this type of apparatus it is intended to emphasize the importance of scoring on all three levels.

In the construction of a PK apparatus no attempt is generally made definitely to exclude mechanical biases from the machine. Instead, the experimenter adopts a rigid alternation of target areas (or target faces in tests where the die face is used as target) so that any biases there might be are cancelled out in the final summary of the results. This method presupposes that any biases the machine might have do not change systematically in the course of the experiment, as the regular alternation of targets is a guard against the influence of bias only if the imperfections do not alter concomitantly. These exploratory experiments did not wholly reach this ideal situation. The bottom tier, which rested loosely on the floor, occasionally received an accidental jolt during the runs. Also, in Series I only, this tier was itself moved every time there was an alternation in target area, to center it immediately below the open (hit) side of the middle tier. In Series II it remained in the same position (see Figure 1) throughout the releases. Because there were 48 target and miss areas it is unlikely that these jolts and changes of position of the bottom tier produced a spurious effect; but the possibility remains, and only experimentation with an improved apparatus can definitely settle it.

In confirmatory research, the manual opening and closing of the trap doors on the top and middle tiers should be replaced by a mechanical system, and all runways made steep enough to prevent the cubes from stopping part way down. The walls of the two upper tiers also should be heightened. It is not the purpose of exploratory testing to provide firm scientific conclusions but rather to point the way to research where such conclusions may be reached. However, there is certainly very strong prima facie evidence for PK in these results. It is difficult to imagine how the scoring, especially on the bottom tier with its 48 alternating hit and miss squares, could have been produced by the manual and mechanical imperfections.

Full insurance against contaminating the results by mechanical biases can only be achieved if the same number of cubes are released for both sides of the apparatus. For instance, if there is a bias in favor of the left side and if also more dice are released when the left side is target than when the right side is, then a spurious positive deviation is the result even if the right side is target as often as the left.

In the present experiment, as in a conventional placement experiment, the number of dice descending into the top tier was always the same. But only those in the hit area of this tier were allowed to continue to the two lower tiers; and as this quantity, of course, varied, unequal numbers of dice reached the two bottom tiers in the different releases. Since it was on the bottom tier that significant results were obtained, it is necessary to examine the data produced on this tier to determine whether a bias favoring one of the areas was associated with a greater number of dice descending on the occasions when these squares were targets.

An analysis of the data shows that 3,133 dice reached the bottom tier when A was target, and 2,838 when B was target. There was therefore a bias in the machine favoring the A targets. However, this was not associated with a suspiciously greater proportion of hits. There were 1,635 hits when A was target, which is 52.20 percent of the total number of cubes which reached that tier on those trials, and 1,469 when B was target, or 51.74 percent of the total. The proportion of hits to the total number released into the bottom tier was about the same with A as target as with B.

The meaningful differences in scoring between the three levels also argue in favor of PK, and there are contained in the distribution effects some further indications that a causal influence extraneous to the mechanical and manual working of the machine was operating.

If, as seems to be the case, PK was responsible for the results, a higher average was obtained with this three-tier apparatus than has so far been accomplished with conventional machines. The average run score of 12.48 on the bottom tier exceeds the highest score thus far achieved, namely Forwald's average of 12.30 given by his deviation of + 500 on 40,000 cube-releases (**4**).

An interesting problem is presented by the negative deviations on the top tier. Since the scores became progressively more positive in the two lower tiers, this negative result is not inconsistent with the experimental hypothesis, but neither is it explained by it. The answer must be sought elsewhere. It may be surmised that subjects preferred the target situation at the bottom tier, and this may have stimulated a matching *negative* attitude toward the top tier. The same effect may have been created here as has been found in so many other psi experiments involving different targets and experimental conditions (e.g., **1, 2, 5**) where the less-preferred situation is associated with psi-missing.

References

1. Cox, W. E. A comparison of spheres and cubes in placement PK tests. *J. Parapsychol.*, 1954, **18**, 234-39.
2. ———. The effect of PK on the placement of falling objects. *J. Parapsychol.*, 1951, **15**, 40-48.
3. Forwald, H. A further study of the PK placement effect. *J. Parapsychol.*, 1952, **16**, 59-67.
4. Rhine, J. B. The Forwald experiments with placement PK. *J. Parapsychol.*, 1951, **15**, 49-56.
5. ———. The problems of psi-missing. *J. Parapsychol.*, 1952, **16**, 90-129.

Box 936
Southern Pines, North Carolina

SOME AVOIDABLE MISCONCEPTIONS IN PARAPSYCHOLOGY[1]

By J. B. Rhine

Any approach to an unknown area of research problems is likely to result in certain misconceptions that are more or less unavoidable. In the early stages of such an inquiry the new descriptive and explanatory concepts are formed partly on the basis of already known principles, and these are destined to prove inapplicable in some degree to the newly discovered territory. The correction of such misconceptions has, of course, only to wait for the more complete exploration of the new ground. Meanwhile, unless invested with disproportionate conviction or pride, they need not seriously interfere with progress.

In parapsychology, however, as is perhaps not unusual, certain quite unnecessary and wholly avoidable misconceptions have acquired some currency, either within or around the active research centers of the field. Since these unsatisfactory ideas must necessarily hamper progress both in research and in its recognition, some effort should be made to correct them. My purpose in calling attention to some of these misconceptions (as I see them) is to initiate discussion that may lead to a better common understanding.

Those which I shall consider in this article have to do with what is actually claimed for parapsychology at this stage—in other words, with just what it is that has been established.

The "Empty Correlations" Concept

The first misconception to be dealt with is best stated in an article in the *American Scientist* entitled "The Present Status of Parapsychology" by Professor E. G. Boring (1): "All you have got yet for extrasensory perception is an observed difference between two frequencies, between hits and misses, and a great deal of ignorance as to what causes the difference." And he adds, "The fact is an empty correlation, one whose novelty could evaporate

[1] This paper was presented *in absentia* at the first convention of the Parapsychological Association in New York City, September 6, 1958.

overnight if its defining conditions came suddenly to be specified more exactly."

It would be more accurate and valid, I think, to say that the "observed difference between two frequencies" in a competently conducted ESP research is part of a systematic experimental project —one involving a complex meaningful design originating with a specific problem. When considered in its actual context, the occurrence of extrasensory perception would be quite inaccurately described as an "empty correlation." If one wished to put the scientific conclusion concerning ESP in its true perspective, he would go back to those puzzling extrasensory experiences that occur spontaneously. He could begin with the collection and analysis of thousands of such cases, the derivation of a testable hypothesis that a human capacity may exist that could, under conditions eliminating all alternatives, produce the effects in question. The satisfactory designing of experiments to eliminate all alternatives and yet allow for the possible operation of the ability investigated would be found to have involved a great deal of systematic planning. Eventually arriving at such experiments as the above author was reviewing, one would see that the observed difference between two frequencies, although a very important element, is only a small part of a large problem-solving operation. Accordingly, to single out this correlation as the sole product of the entire study ("all you have got yet") does not even offer a good caricature of the research. The logic of this "empty correlation" argument is to imply that because one does not know enough about a research finding to avoid having to use statistics one, therefore, really knows nothing. While it is well to guard against claiming too much regarding the findings of a research, it would be to misconceive the entire mission and meaning of science to misrepresent them in this manner. The above characterization then, of what has been established in parapsychology is both inadequate and misleading.

The Concept of ESP as an Established Species Character

Another misconception as to what parapsychologists have claimed is represented by a critique of the case for ESP given by the late Dr. R. Von Mises (8) at the joint meeting in December, 1949, of the Philosophy of Science Association, the American Philo-

sophical Association, and Section L of the American Association for the Advancement of Science. His criticism was directed against the claim that ESP has been established as an enduring capacity, first, of the human individual and, second, of the species.

On both these points, however, the attack was aimed at a straw man. It is quite true that parapsychologists naturally suppose when an individual gives evidence of ESP that the potential ability was there *before* the test and will remain *after*. However, no one had claimed to have proved such a generalization, and accordingly no one needs to defend it. Likewise, after having given ESP tests to a wide range of different individuals and having found evidence of ESP in a large number of cases, and at the same time having come upon various attitudes, test conditions, and types of approach that appeared greatly to affect success in ESP, workers came to suspect that the capacity is indeed a general one and demonstration a matter of finding favorable conditions. That this conception, too, was only hypothetical and had not been presented as established was well known to all who had given extended attention to the field. The distinguished scientist, however, even though he was examining the subject for a learned society, had projected into the findings of parapsychology a misconception that was not warranted by the actual wording of the research reports.

The Concept of Production at Will

A third misunderstanding that has proved much more extensively misleading than the two mentioned above is the view that psi has been set forth as reproducible under specifiable conditions and that it ought, therefore, to be subject to the challenge of demonstration and the demands of application. "Why don't you break the bank at Monte Carlo?" one challenger will say, while another will misconceive ESP as something around which to build a television or radio show. Critics will maintain that if psi really occurs, then it should be expected to occur on call. This third misconception has bothered even some of those who have taken an active and serious interest in this field by leading to premature expectations and consequent disappointment with results.

The most widely circulated publicity was given to this erroneous view by Dr. G. R. Price in 1955 with his demand that an ESP test

be conducted under the supervision of a jury of a dozen skeptical scientists following a procedure which he outlined in his published article (**5**). The logic of the challenge was that if such an experiment were made and failed to show significant evidence, it would presumably refute the conclusions of those many parapsychologists who had evidence of ESP. The rest of the argument was that unless such an experiment were carried out and the results were positive, the case for ESP could not be taken seriously.

Back of this sort of challenge, in all its many and various forms, is a fallacious notion of what the ESP findings really are and have been stated to be. The author's mistake was that he tried to pin parapsychology down at a stage it has not yet reached and has not been claimed to have reached. It is easy to see the misconception involved here if we recall that there are two distinct stages in this investigation. In the first of these, the *occurrence* of the phenomenon is in question. The existence of ESP has now been established and the first stage of parapsychology has been successfully completed. The objective of the second stage is to bring the phenomenon to the point of production on demand, repeatedly and under specifiable conditions. But while much progress has been made, this second goal has not as yet been reached; and I know of no worker who claims that it has. That is the status of parapsychology today.

Of course, in the sciences many phenomena and effects have gone through comparable stages. Lightning, for example, and immunization were known to have occurred long before they could be reproduced on demand. The distinction is simple enough when attention is given it. The investigators who first approached psi phenomena rightly asked first whether, under conditions that safely excluded alternatives, an extrasensorial mode of perceiving could be demonstrated to occur at all. When this initial query established the fact that *ESP does occur,* it was necessary at that stage only to shield the hypothesis under test from all reasonably conceivable alternatives. The results under these conditions showed that ESP *had* occurred. Proving, thus, that ESP could occur was a clear advance in itself, and naturally led to the next step, the search for controlled occurrence.

No one, of course, has been any more interested in the question of how to induce repeated functioning of psi than the investigators who first gained reliable evidence of it. In fact, the many exploratory activities that followed its demonstration were aimed at finding factors and conditions that would influence reproduction of the phenomenon and bring its recurrence under control. So clear and manifest have these directions of effort been that only those who are but superficially acquainted with the progressive activities in parapsychology could miss the point that it is largely the search for means of controlling psi (of course, through increased knowledge of it), that has dominated the research policies of many workers for the past twenty years. And it is quite in line with the history of science to anticipate that in time the production of ESP effects on demand, and even its reliable application, may be expected to follow.

Probably no one, on cool reflection, would seriously adhere to this third misconception once it has been sharply pointed up. The alternative to which some will turn, however, is to reject any finding that cannot be demonstrated at will. But this is tantamount to saying that nothing can be scientifically established at all until it can be reproduced on demand. It would be to say that all those scientific undertakings in which one has to wait for the phenomenon to occur and then capture it by carefully staged precautionary measures must be discarded as useless until, if ever, the occurrences become controllable or regularly recurrent. Fortunately, such a misconception does not necessarily represent science merely by being given prominence in *Science*.

Consequences of the Third Misconception

These three misconceptions have been discussed as they have been represented to the public in critical appraisals of parapsychology. But I indicated also that more serious consequences than criticism follow, especially from the last-mentioned of these misconceptions, consequences which are manifested especially by some of those who have undertaken to conduct ESP experiments and have failed. I use the word "failed" because at the present stage of parapsychology (in which no one can claim to know how to produce ESP on demand) it should now be clear that an ESP experiment could hardly be considered successful if it produced no evidence of

ESP. There are at this stage two main requirements for a good ESP experiment: first, the experimenter must be able to induce the phenomena to occur, and second, he must be able to control the conditions so as to rule out alternative possibilities. The two requirements must both be met to make the experiment worth while. Without ESP it cannot be an ESP experiment and no one can rightly conclude anything from it.

An example of how this erroneous conception of what is claimed for ESP can mislead people engaged in actual experimentation is rather well illustrated in an article published in 1954 in *Science* by Smith and Canon (**7**). In their experiment the design was drawn up without much regard for the psychological conditions required for such testing. The logic of the experimenters seemed to be that if ESP does occur in human beings, as parapsychologists claim, it should be expected to occur in their experiment as they designed it and conducted it. But it did not occur. They considered, then, that the failure threw doubt on the validity of the original claims; and since the editor of *Science* gave space to the publication of the otherwise pointless article, he evidently concurred in the fallacy of the authors.

But in the very nature of the case, the insignificant result of this experiment could not prove anything; and its publication only shows how far the authors were led astray by their misconception of the actual status of ESP results. It will clarify this point to focus more sharply on the real issue; even had the procedure in this case been an exact duplicate of, let us say, the Pratt and Woodruff series or the Soal and Goldney work (and it was very far from that), the logical fallacy in question would have been the same. Even, to make the issue an extreme case, had Pratt or Soal or a half-dozen other leading parapsychologists claimed that they *themselves* could count on getting again significant ESP results with their own subjects under the same conditions—and I know of no one who would make such a boast—the logic of the Smith and Canon experiments would still have been as faulty as ever. For at this stage in parapsychology no second attempt can be counted on as a precise and dependable replication of an original experiment. With all the many uncontrolled factors and variables of human personality involved in such an attempted repetition, no failure can have any

generalization value whatsoever. It is time for those who appear to think differently to present a logical case for their views instead of merely producing more of these pointless failures.

Confirmation Instead of Repeatability

Perhaps nothing in the entire conceptual structure of parapsychology is more in need of clarification than the distinction between what is indefinitely called the problem of *repeatability* and the very real need for *confirmation* of findings—a basic requirement for all research fields. The repetition of an experiment in one of the more exact sciences, where pure substances and accurately controlled conditions are possible, is an important method of testing the truth of a claim. When a new finding in such a field is made public, a cautious student will wisely ask: "Has it been repeated?" By that he means to ask if substantially the same experiment has been conducted by another worker and with similar results. This is the customary way of seeking confirmation in research areas in which the method is appropriate. But repetition in that sense is not a universal way of confirming a discovery. In a field (for example, geology) in which experimentation is not a possible method of inquiry, the goal of confirmation is achieved by independent repetition of observations. In any area of science where wholly spontaneous occurrences of nature are dealt with, there is likely to be a stage during which confirmation can only be made by purely non-experimental, descriptive methods—as, for example, in the early study of meteorites, or of a new disease.

But when, as in parapsychology, inquiry has advanced to the stage at which experiments can be designed to eliminate all alternative hypotheses and still allow the hypothetical phenomenon to occur, appropriate methods of confirming are needed, too, in keeping with the advance, but in keeping also with its limitations. Accordingly, even though the controlled repetition of a psi phenomenon is still uncertain because of the semi-spontaneous nature of its occurrence, adequate confirmation has taken place repeatedly. It does not, of course, matter how many attempts were made under conditions lacking something (for example, a productive subject) essential to the effect in question. Such fruitless attempts cannot disprove or in any way lessen the value of those which have more successfully

caught the necessary factors and conditions. As it is, confirmation of ESP has been going on cumulatively for decades, far and away ahead of the controlled performance for which we still are waiting. But the fact is, repeatability-on-demand must be distinguished as really another problem, and it must await its turn for investigation—more of a wait than some had anticipated.

Parapsychology, as the literature shows, has had to rely for its confirmations upon investigators who have been able either to find certain subjects already endowed with capacities and attitudes conducive to successful performance under test conditions, or else to find experimenters able to give their subjects the requisite motivation and orientation to perform well under such controls. This is nonetheless confirmation of the claims that have actually been made; and for anything that has not been claimed, it is pointless to expect to find supporting duplication.

Establishment by Majority Rule?

Not only does this misapplication of the test of repeatability to the case for ESP underlie the experimental proposal of G. R. Price and the experiment referred to as actually carried out by the unsuccessful experimenters mentioned previously, but a third and rather more extreme application of the misunderstanding is represented in an article to appear in the *American Psychologist* by Dr. J. C. Crumbaugh (2). In this paper, after stressing repeatability as a standard (and presumably unexceptional) requirement for science, the author proposes still another experimental program for parapsychologists, an even more elaborate one than that of Price. In his plan a large number of experiments are to be conducted, each by a different worker but with a common design and procedure. The case for ESP is to rest on whether or not a majority of the different experimenters obtain independently significant evidence of ESP. Here we at least have something new proposed for scientific method.

Important as a majority vote may be in some situations, however, it is not a satisfactory criterion of truth in science, nor even a measure of significance. But this feature is after all rather incidental. The essential misconception in this proposal lies in the assumption that the claims of parapsychology have progressed so far that results in a given case or percentage of cases ought to be

reliably expected and predicted. In other words, the investigator in this case would be testing, not the proper hypothesis that ESP occurs, but the as-yet-unclaimed idea that a majority of any group of experimenters with any group of subjects should be expected to obtain significant evidence of ESP by merely following a defined procedure. I need not repeat that no one at this stage vouches for results with such a "defined procedure." This is misconception-number-three again, pure and simple. Add on the difficulty that even if ESP *were* ready for such a conception, it is hard to see what conclusive difference it would make in the general case for that capacity whether the number of experimenters who obtained significant results were 51 or 49 per cent. There is no such criterion of scientific truth.

Numbers of proper confirmations, however, do mean something, even if not in terms of percentages. Some will even ask how many are required to establish the case for ESP. That is a matter of personal judgment, and one that depends, too, upon the individual confirmations. The point stressed here is that the number of non-significant researches has no bearing whatever on the general case for ESP. Far too many ways of failing to provide the necessary conditions exist for chance results to have reliable meaning of any kind.

So this proposal, like Price's challenge and the Smith and Canon logic that "our failure discounts your success" can be listed as due to a mistaken conception of what is known and claimed in parapsychology.

The Pooling Fallacy

There is still a fourth consequence of this misconception of the reproducibility of psi on demand—one that has been a source of concern even to a few of the productive workers within the field. Accordingly, if the present discussion succeeds in clarifying only this one point of confusion, it will have accomplished a mission of some value.

The form which this misunderstanding usually takes is an insistence that all the known experimental results bearing on a given problem in parapsychology must be pooled and treated in their totality in order to avoid improper statistical selection, an idea that

was first given expression in print by Dr. C. Leuba (3) in 1938. This demand is aimed at avoiding the improper selection of favorable collections of data from what might be only a chance series of test results. The publication of any significant research results not accompanied by all other work done on the problem by others everywhere is claimed by proponents of this view to give an unfairly favorable picture of the general case for the type of psi phenomenon under consideration. The implication is that no unit of work relying on statistics, no matter how independently conducted or how well designed, can be considered entirely by itself—on its own merits. If somebody else (or many others) should do a related experiment, the two (or all) studies must then be considered together; otherwise there is the risk that one would publish a report because his results were positive, while those who obtained negative results would not consider their investigation worth reporting. This, it is believed, could only result in the error of improperly selected data. But if, according to this viewpoint, the only correct way is to publish all the work done on the problem in one complete summing-up, the investigation of such a problem as that of ESP would be hopeless. How could a research worker ever know where he stood until all reports were in? And how could unpublished work be rounded up? Probably no one would insist on carrying pooling so far, but this absurd extreme only serves to point up the fallacy that runs through the whole idea of compulsory pooling.

It should be pointed out that this exorbitant demand for pooling is, if it belongs anywhere, primarily a statistical issue; but as far as I know, it has not been urged or even suggested by statisticians. As a matter of fact, it is fairly obvious that such a general requirement would go far to render statistical methods valueless; it would raise a serious question as to the integrity and sufficiency of any single experiment, however complete in itself. In short, if every experiment had to be pooled with all similar ones, published or not, before any final evaluation would be safe, there would be no statistical methods. Here again, the adequacy of confirmation is the point to emphasize, but it is not to be confused with this fallacy of a need for a universal combination of data of similar type.

It is true, no one could properly approve the presentation of an *incomplete* experiment in which a part of the experimental data was

selected without due correction. But that is quite a different matter and, though important in itself, is not involved in this issue. I am considering only the methodological independence of complete, separate, and well-designed research units.

Nor do I mean to imply all combinations are worthless. Surveys involving the pooling or combination of the work of a given laboratory or even of a larger selection have been presented from time to time in the literature of parapsychology and may continue to be. For example, in *Extrasensory Perception After Sixty Years* (6), the authors undertook to summarize combinable experimental results secured by many different investigators. This was not a statistical necessity; it was rather an exercise to give added assurance to the student of this field at a stage when the ordinary logical and experimental criteria of science were not enough and were not well enough appreciated to allay the critical apprehensions of a great many conventional scientists in their appraisal of parapsychology.

There are, however, really two errors involved in this insistence on the necessity of combining all comparable psi research results: The first, already discussed, is the statistical illusion that no single psi experiment can in itself be so completely adequate as to justify its conclusions regardless of what other experiments along similar lines may be conducted. The second fallacy is another consequence of misconception-number-three, the mistaken notion that parapsychologists claim that psi may now be considered producible on demand. It assumes wrongly that enough is known to conduct two or more experiments so nearly alike as to have a right to expect like results. Only then is pooling a proper and profitable procedure—that is, only when there is no difference in the conditions and factors that could possibly affect the results. With the great complex of variables still existing in all psi experiments, there is no sound basis to justify this proposal of throwing the work of all experimenters, all subjects, and all conditions into one big evaluative hopper. The argument, in summary, is that psi experiments could not be expected to have enough in common to allow one to negate or alter the value of another, as would occur in the pooling under consideration. Negative results, it should be remembered, are always meaningless whenever too little is known about the phenomenon to allow reliable explanation of the factors responsible for

failures. Only when the conditions essential to psi are well enough known to make it reproducible on prescription, and hence to duplicate the essential conditions from one test to another, can this urge to pool results be granted. By that time, however, it would have no point.

Perhaps an example may help to make these remarks more clear. I will take for illustration a report by Mangan presenting evidence for precognition in tests with a selected subject (**4**). The primary respect in which the Mangan experiment advanced beyond the point reached in previously reported tests of the precognition hypothesis was in the use of a more complicated method of selecting the random order of targets. As the report states, this method (which imposed a complicated machine calculation upon digits obtained from throwing dice as a means of entering a random number table) was suggested a few years previously in the Parapsychology Laboratory. Before the Mangan experiment was conducted, different members of the research staff had carried out tests with a variety of objectives, some trying out certain procedures, others modifying the test conditions, and all working with a large number of subjects and their various attitudes and abilities. These projects were exploratory in that none were set up on the basis of significant results in previous experiments. In the case of the Mangan experiment, the work was done with a subject who had already given significant results in an ESP experiment, and the precognition test was planned on the basis of the earlier work to see if this particular subject could also respond to future targets selected in the manner indicated. Could anyone seriously have thought while the experimenter was still planning his work that he could not answer his question unless he intended to pool his results with those of all previous (or later) precognition tests using a similar method of target selection? It is obvious, on the face of the matter, that any amount of insignificant data from precognition tests by other experimenters with other subjects could have no bearing upon the question of whether *this* experimenter working with *this* subject under the psychological conditions of *this particular investigation* could obtain significant results. Thus the design of the Mangan experiment not only did not require but also did not allow such pooling of results.

We would do well to keep reminding ourselves of just what we *do* claim and *should* expect. We must, in the present instance, clearly distinguish between the need of confirmation and this pooling fallacy, just as we must do between confirmation and repeatability. We must remember that we are still groping in an exploratory way toward the knowledge that is needed for more control over performance; and while this stage lasts, we must carefully avoid the trap into which the misconception under discussion would put us—the implication that we can depend on and control the recurrence of psi effects in a given situation. As long as the parapsychologist does not himself make such a claim, he need not answer for it. But he has also to guard against erroneous conceptions in the expectations of others as well as against easy generalizations in his own thinking.

* * *

In summary, it may be said first that the investigations of psi phenomena have achieved a great deal more constructive scientific progress toward a rationally integrated body of knowledge than just "empty correlations." In fact, a vital new branch of psychology has been initiated. Secondly, although the hypothesis that psi is an enduring function of the individual and of the species is entirely reasonable, it is unprovable at this stage and must be considered only as a working hypothesis. Third, while the occurrence of psi has been established as a clearly distinguished and definable effect, it has not yet become an effect reproducible on demand under prescribed conditions. The validated claim is that the occurrence of psi as a natural phenomenon has been established by virtue of having been repeatedly confirmed in many careful experiments. Moreover, much has already been learned by the repeated occurrence of psi effects under well-controlled conditions of observation that should be helpful in bringing the ability under better control. This is an early stage of parapsychology which must not be allowed to appear as more final than it really is or the effect would be misleading. We who work in this field should not allow ourselves even to consider as yet challenges and invitations predicated on the assumption that psi is in any degree reproducible on demand; nor should we indulge in combinations or other treatment of results through unwarranted and premature assumptions of the degree of

common character and lawfulness in our experiments which that would require.

References

1. Boring, E. G. The present status of parapsychology. *Amer. Scientist,* 1955, **43**, 108-17.
2. Crumbaugh, J. C. ESP and flying saucers. [Awaiting publication in the *Amer. Psychologist.* See also comment to follow.]
3. Leuba, C. An experiment to test the role of chance in ESP research. *J. Parapsychol.,* 1938, **2**, 217-21.
4. Mangan, G. L. Evidence of displacement in a precognition test. *J. Parapsychol.,* 1955, **19**, 35-44.
5. Price, G. R. "Science and the Supernatural," *Science,* 1955, **122**, 359-66. See also articles in reply in *Science,* 1956, **123**, 7-19.
6. Rhine, J. B., Pratt, J. G., Stuart, C. E., Smith, B. M., and Greenwood, J. A. *Extrasensory Perception after Sixty Years.* New York: Henry Holt, 1940.
7. Smith, K., and Canon, J. *Science,* 1954, **120**, 148. See also comment on this article in *Science,* 1955, **121**, 808.
8. *Parapsychology Bulletin,* No. 17, Feb., 1950.

Parapsychology Laboratory
Duke University
Durham, North Carolina

THREE NEW ESP TEST MACHINES AND SOME PRELIMINARY RESULTS[1]

By W. C. STEWART

ABSTRACT: This report describes an exploratory ESP experiment in which three ESP machines were used. Each of them provided a random target series, a mechanism for making the guesses, and counters which tallied the number of trials and the number of hits. These three machines were used in a competition in which university students were invited to participate during two days of testing. As a stimulus to high scoring, cash prizes were awarded to the subjects having the best results. There were 99 subjects, several of whom had marked deviations in the ESP tests. The total scores on two of the machines were insignificant. The third machine, constructed and operated by the author of this report, gave statistically significant evidence of ESP.

Mr. Stewart was in his senior year in the College of Engineering, Duke University, when the experiment was carried out.—Ed.

INTRODUCTION

THE NEED for machines capable of testing for ESP as well as for PK has been recognized for years. The busy parapsychologist often finds that shuffling cards, keeping scores, recording random digits, and closely observing the subject in his extensive experiments is a difficult task. Obviously, a machine which would reliably provide a random target, record data accurately, and be foolproof in its operation would relieve the experimenter of several monotonous chores and allow him more flexibility in creating a desirable environment for his experiment. Though there have been some successful ventures with mechanical test devices (**1**, **2**) many ingenious ESP machines that have been built in the past have failed from mechanical difficulties or because the inventor was unable to pursue the necessary further development of his idea.

In 1956, Dr. J. L. Artley of the electrical engineering department of Duke University produced an ESP test device with a cylinder whose rotating motion was very nearly random. In using this device, the subject made his responses by pressing a button whenever it was his impression that a marked portion of the cylin-

[1] This research was first presented during the Convention of the Parapsychological Association on Sept. 5, 1958.

der was facing a given direction. In the summer of 1957, while working with the Duke Parapsychology Laboratory, Dr. Artley modified his machine so that the target, now a circular disc, could be stopped in one of five randomly selected positions corresponding to the five standard ESP symbols. The results of his work indicated that a machine of this nature might be suitable for the purpose of testing ESP. Encouraged by the interest of several students, Dr. Artley and Mr. D. B. Young, Instructor in Engineering at Duke, introduced a research and development project in ESP instrumentation available as an elective to electrical engineering seniors. In this project Messrs. J. E. Jenkins, G. E. Mott, and W. C. Stewart each during 1957-58 designed and constructed a machine.[2]

Apparatus

Mr. Jenkins' apparatus employs a tape for producing a random target sequence in which punched holes are spaced at intervals given by digits from a random number table. The tape is wound on two reels, similar to a typewriter ribbon, and placed so that it moves behind a small window equipped with a shutter. The subject turns a dial, winding the tape from one reel to the other, until he has the impression that a hole in the tape is behind the closed window. Depressing an electrical switch records his try and a hit, if one is made, on two digital counters. The subject then turns the dial until he makes another guess by working the switch. Electrical circuit operation is such that only one guess can be made at any single tape position. The probability of obtaining a hit with this device is 1 in 4.

Mr. Mott's machine contains 12 small lamps which may be lighted one at a time in sequence. A geiger counter provides a randomly varying number of impulses for periods of time on the order of one second, and is used to select which lamp will be lighted for each try. After the lamps are covered, the subject attempts to guess by ESP which one is lighted. He turns a pointer towards that lamp and depresses an electrical switch. Digital counters record the number of guesses and hits, with a new target being auto-

[2] The author will be glad to supply a more detailed description of any of the machines to interested readers.

matically selected after each try. Probability may be varied to values of 1 in 2, 3, 4, 6, or 12.

Stewart's ESP apparatus is similar to Mr. Mott's except that it has only five lamps, giving a probability of 1 in 5, and uses a different type of random target generator. Contained in the circuit is a time-interval switch with an accuracy to two significant digits. Random factors affecting its mechanism cause higher order digits to occur in random sequence. An electrical counter times the operation of the switch to the fourth significant figure and lights one of the five lamps, each of which corresponds to two of the 10 possible digits that may occur in the fourth significant figure. Extensive tests for randomness, suggested by Dr. J. G. Pratt, were applied to the digit sequences generated in this manner and gave proof of a sufficient degree of randomness. In testing, the subject rotates the array of lamps by means of a dial until the one he believes to be lighted is beside an indicator. Pressing a button records the guess and hit, if one is made, then selects a new target for the next try.

Procedure

The three ESP test devices were first demonstrated at the Duke Engineering Show in March, 1958. Visitors were invited to operate the machines, and records of their scores were kept. One hundred thirty-one persons participated in these informal tests, doing an average of 43 tries per person. Another smaller demonstration was carried out in April during a student convention on the Duke campus. The results of these two demonstrations were promising, especially when the subjects were allowed to know their number of hits during the test. It was therefore decided to carry out another experiment under stricter conditions. This research was conducted on May 10 and 12, also on the Duke Campus and with members of the University community as volunteers. It was decided beforehand to have each subject make 50 trials on each machine and to let the number of participants be determined by the number who could be accommodated during the two days of testing.

Each of the machines was checked periodically in order to ascertain that it was operating properly. In one instance, it was found that a malfunction had occurred in Stewart's device. The block of

data between that point and the previous check point is not included in the assessment of the results.

The experiment was conducted in the form of a contest in which cash prizes were awarded to the highest scorers. During the two days of the contest 99 subjects were tested on all three machines. An additional 16 subjects could not wait for their turn on all three machines but had to leave when they had been tested on only one or two. Miss Margaret Anderson of the Duke Parapsychology Laboratory assisted in the experiments and kept a record of all scores. Each of the three operators made an independent record of the scores made on his machine, and this was checked against Miss Anderson's record. No discrepancies were found.

Results and Discussion

The results are shown in the following table. The over-all deviations on Mott's and Jenkins' machines were not significant but the total on Stewart's was, with a CR of 3.14, even with the appropriate statistical correction ($3P = .0025$). Also, several of the scores by individual subjects are impressive, though account must be taken of the many chance scores in assessing the significance of the relatively few individual results which stood out.

Summary of Data from ESP Contest, May, 1958

	Jenkins Machine	Mott Machine	Stewart Machine
P-value	1/4	1/12	1/5
Total subjects tested	115	111	75
Total trials	5750	5550	3750
Expected hits	1437.5	462.5	750
Total hits	1368	448	827
Deviation from chance	−69.5	−14.5	+77
Over-all CR	(−) 1.05	(−) 0.68	(+) 3.14
Largest individual CR	(+) 2.94	(+) 2.97	(+) 3.53
Number of individual CR's greater than 2.32			
a) positive	1	1	3
b) negative	1	0	0

While statistical analyses are enlightening, perhaps a few clinical observations would be in order. The encouragingly successful outcome of the tests is likely to have been due to the atmosphere of informal, friendly competition both between the subjects and be-

tween the experimenters, each of whom was eager to demonstrate the virtues of his machine. It was noticed that scores tended to be higher in general when a group of persons was present who manifested an open, friendly, competitive attitude; but the highest individual scores did not occur at these times. The presence of seemingly distracting noises or activities around the subjects apparently had no influence on scoring trends. Although the contest ran for a total of 11 hours, the three prize-winning scores were obtained in a single one-hour period, namely, from 2:00 to 3:00 p.m. on May 12.

It may be stated in conclusion that mechanized testing promises to be a significant aid in experiments when it is desired to have an informal situation which nevertheless encompasses all scientific safeguards.

In such tests the random presentation of targets, as well as the scoring, is accomplished by the device, so that the experimenter can devote himself to the psychological aspects of the tests. However, the devices herein described are in an early developmental stage, and much time and extensive effort must be expended in order that their full research potential be realized.

References

1. Tyrrell, G. N. M. Further research in ESP. *Proc. Soc. psych. Res.*, London, 1937, **44**, 99-168.
2. Webster, D. D. An automatic testing and recording device for experiments in ESP. *J. Parapsychol.*, 1949, **13**, 107-17.

School of Engineering
Duke University
Durham, North Carolina

TERMINAL SALIENCE WITH MULTIPLE DIGIT TARGETS[1]

By Carroll B. Nash and Martin G. Durkin

ABSTRACT: This paper reports an exploratory research on the question of whether the subject in an ESP experiment responds to the target as a unit (in accordance with principles of Gestalt psychology) or whether he responds to the various components of the target separately. The targets in this experiment consisted of numbers having from one to five digits apiece. There were 45 subjects, who completed a total of 1200 runs. They showed no tendency to score on the total target, though there was statistically significant evidence of ESP in the distribution of hits on the separate digits. The subjects showed a marked tendency to score on the outer digits of the numbers and to miss the inner ones, a fact which the authors interpret as evidence of a Gestalt effect.

Dr. Nash is well known to Journal readers. Mr. Durkin, at the time of the experiment, was an undergraduate student in Biology at St. Joseph's College.—Ed.

Introduction

The targets to which subjects are asked to respond in ESP tests may have a single aspect (such as an ESP symbol) or more than one aspect (as a colored ESP symbol). In a review of published research with multiple-aspect material, Foster (1) concluded that in some of the experiments the subject responded to the complete target as a unit. In other words, he tended to guess all the aspects of the target rather than only one. If this is typical of ESP responses, it would indicate that there is a Gestalt effect in ESP. However, there are also experiments in which the subjects seemed to respond to only one aspect, and further research on this question is necessary before it can be determined whether or not ESP subjects generally respond unitarily to all the parts of composite targets. The exploratory research reported here was designed to throw light on this problem. In this experiment, where some of the targets consisted of multiple-digit numbers, the validity of the Gestalt hypothesis would be indicated if there was a tendency to score on all the numbers of the target rather than on only one.

[1] The research was supported by a grant from the Parapsychology Foundation, Inc. It was first reported at the Convention of the Parapsychological Association on September 4, 1958.

Procedure

The experiment was performed in 1955-57. As planned, it consisted of 100 sets. Fifty-five subjects were tested, each of whom completed from one to five sets apiece. The set consisted of one of each of five kinds of run, namely runs with one-, two-, three-, four-, and five-digit targets. The run was made up of 20 trials regardless of the number of digits of which the targets were composed. Hence each set comprised 300 single digit trials. Within the set, each of the five kinds of run occurred in the five chronological positions an equal number of times in the whole experiment. The targets consisted of the digits 0-9 in random order. They were transcribed from a table of random numbers (3) to record sheets, each containing five runs, by the junior author, who was also the experimenter.

The targets were covered during the test by an opaque piece of paper and placed in front of the experimenter, who was screened from the subject. The subjects were tested singly and completed one set per session. They made their calls orally and the experimenter recorded them in a column next to the column of concealed targets. Calls that agreed with target digits in the corresponding positions were counted as hits. The number of possible hits per target is the same as the number of digits in the target. The results were given to the subject at the conclusion of the session.

Forty-three of the subjects completed the Guilford-Zimmerman Temperament Survey (2) which measures on an eleven-point scale the personality traits of general activity, restraint, ascendance, sociability, emotional stability, objectivity, friendliness, thoughtfulness, personal relations, and masculinity. The results of the experiment were independently checked.

In four of the 100 sets, the experimenter was also the subject. His total deviation was +1 and as his results are not noticeably different from the rest they are considered jointly with it.

Two subjects carried out one set under hypnosis with positive suggestion and one set in the waking state.

Results

The results obtained with each of the five kinds of target are presented in Table 1. The total deviation of − 42 in 30,000 trials

Table 1

TOTAL SCORES AND SCORES ON DIFFERENT DIGITS

No. of Digits in Target	No. of Trials	Deviations	Deviation for Each Digit				
			Digit position within the target*				
			1st	2nd	3rd	4th	5th
1	2,000	+ 4	+ 4				
2	4,000	−39	−29	−10			
3	6,000	−24	−12	−21	+ 9		
4	8,000	+ 4	+22	−12	−18	+12	
5	10,000	+13	+12	−14	−13	− 2	+30
Total	30,000	−42					

Total deviation on inner digits: −80
Total deviation on outer digits: +73
CR diff. = 3.29
P = .001
No. subjects with higher deviations on outer digits: 36
No. subjects with higher deviations on inner digits: 15
$\chi^2 = 8.65$ (1 d.f.)
P = .0033

* The expected score is 200 for each digit position.

is not significant. There is no indication of a differential favoring multiple hits per target.

The deviation of hits in each digit position is also given. The total deviation of the outer digits (the first and last digits of the three-, four- and five-digit targets) is + 73, and the total deviation of the inner digits is − 80. The CR of the difference is 3.29 with a P-value of .001.

This terminal salience was widespread among the subjects; the score was higher on the outer than on the inner digits with 36

Table 2

ESP-PERSONALITY CORRELATIONS IN PRESENT AND PREVIOUS RESEARCH

Personality Trait	Nash-Durkin	Nash-Nash
General Activity	0.31	0.36
Restraint	−0.04	−0.23
Ascendance	0.14	−0.01
Sociability	0.40	0.28
Emotional Stability	0.24	0.26
Objectivity	0.21	0.25
Friendliness	0.20	0.21
Thoughtfulness	−0.07	−0.28
Personal Relations	0.10	0.03
Masculinity	0.27	−0.09

subjects and the reverse with only 15. The chi-square is 8.65 with a P-value (1 d.f.) of .0033.

The correlations between ESP scores and scores of personality traits of the 15 subjects who performed three or more sets are presented in Table 2. The ESP scores of subjects with fewer than three sets are believed to be less reliable. For purposes of comparison, the correlations obtained in a similar experiment with 24 subjects (**4**) are listed. None of the correlations in the present experiment is independently significant. The combined scores of the two hypnotized subjects were + 7 in the waking state and — 12 under hypnosis. The CR of the difference is 2.59 with a P of .0047.

Discussion

The main purpose of these experiments was to ascertain whether or not the subjects tended to respond to the target as a unit. The results contain no indication that they did. However, as Dr. Gardner Murphy has pointed out,[2] the higher score on the outer digits represents a structuring in the ESP responses which, in point of fact, is wholly consistent with Gestalt concepts. It thus seems that, although the subjects did not respond unitarily to the targets, they nevertheless showed evidence of a Gestalt effect in the distribution of their scores on the digits of which the targets were made up. This type of distribution is similar to the terminal salience and U-curves which have occurred within the divisions of many other psi experiments. Terminal salience has been found in score segments as short as five trials in length (**6**). While this type of distribution effect was not predicted in the present experiment, it was the intention from the beginning to determine the deviation in each digit position. In view of this and the highly significant CR of the difference between the total scores on inner and outer digits, as well as the significant consistency with which the individual subjects followed this trend, it seems safe to conclude that the salience effect was not an artifact due to indiscriminate *ex post facto* statistical analyses.

Since the targets were recorded considerably in advance of their use and were covered during the run, clairvoyance is considered a

[2] During the discussion following the presentation of this research at the Convention of the Parapsychological Association.

more likely hypothesis than telepathy, although the latter was not precluded.

Nothing can be concluded regarding the function of hypnosis from tests with only two subjects except that there was evidence of ESP in the marked difference in scoring between the trials in the hypnotic and waking states.

With regard to the personality correlations, though these are not significant, they are in agreement with the results of comparable previous experimentation (**4, 5**). The correlations in the present experiment are of the same sign as those obtained previously (see Table 2) with the exception of the correlations between ESP and the traits of ascendance and masculinity which were considered in the previous experiment to be too low to indicate a relationship. On the basis of their experiments, Nicol and Humphrey (**5**) gave an expected range of correlation between ESP and the personality traits of restraint, emotional stability, and thoughtfulness. The correlation of each of these traits falls within the range they predicted.

References

1. Foster, E. B. Multiple aspect targets in tests of ESP. *J. Parapsychol.*, 1952, **16**, 11-22.
2. Guilford-Zimmerman Temperament Survey. Sheridan Supply Co., Beverly Hills, California, 1949.
3. Kendall, M. G., and Smith, B. B. Tables of Random Sampling Numbers, Tracts for Computers XXIV. London: Cambridge University Press, 1939.
4. Nash, C. B., and Nash, C. S. Checking success and the relationship of personality traits to ESP. *J. Amer. Soc. psych. Res.*, 1958, **52**, 98-107.
5. Nicol, J. F., and Humphrey, B. M. The repeatability problem in ESP-personality research. *J. Amer. Soc. psych. Res.*, 1955, **49**, 125-56.
6. Rhine, J. B. Terminal salience in ESP performance. *J. Parapsychol.*, 1941, **5**, 183-244.

Parapsychology Laboratory
St. Joseph's College
Philadelphia 31, Pa.

SOME COMMENTS ON THE CONTROVERSY REGARDING DOWSING RODS[1]

By Carl von Klinckowstroem

[Editorial Introduction: The practice of dowsing presents a problem that has challenged the interest of parapsychologists since the beginning of systematic research in our field down to the present time. In England and America scientific thinking has been focused primarily upon the question of whether the practices of the dowser—if his claims for successful location of underground water or other substances have any basis in fact—provide an instance of the practical application of ESP through unconscious muscular responses.

In spite of the long history of interest in dowsing, the problem remains in need of further research. The field situations under which the practitioners of the art are accustomed to work do not lend themselves to the kind of controlled, quantitative test that is needed to check the dowser's ability, and hardly enough has been done under satisfactory "artificial" experimental conditions to warrant final conclusion. From the few tests that have been reported, two generalizations are justified. One is that it is very difficult to motivate the dowser in the objective test situation, where his success or failure can be strictly ascertained, to a degree that compares with the strong motivation or self-assurance he seems to feel under actual field conditions. The second is that dowsers have nevertheless obtained results in several quantitative experiments[2] which lend support to their claims, even though the performance has always been far below the level of almost complete success which is frequently claimed for the practice.

The following paper by a German scholar who has long been interested in the problem of dowsing is one that the editors thought worth making available to American and English readers. Count von Klinckowstroem has made an informative review of the scientific researches and debates on the problem in Europe in recent years, particularly in Germany, Holland, and the Scandinavian countries. It appears from this survey that the amount of serious attention devoted to the questions

[1] This is a slightly abridged and revised edition of a paper which originally appeared in *Die Wasserwirtschaft*, 1958, **48**, 237-39. The translation is by Mr. J. Quittner.

[2] Barrett, W. F., and Bestermann, T. *The Divining Rod.* London: Methuen Co., 1926.

Rhine, J. B. Some exploratory tests with dowsing. *J. Parapsychol.*, 1950, **14**, 278-86.

Two reports on experiments with dowsing. *Parapsychology Bulletin*, Dorothy H. Pope, Ed. No. 20, Nov. 1950.

raised by dowsing has been much greater in those countries than it has been during the same period in the English-speaking world and in America in particular.

But the publication of this paper also serves a second purpose, and this is to call attention to the fact that the author of this review as well as those whose works he cites approach dowsing from the point of view of hypotheses which would, if supported by sufficient evidence, remove the practice from the field of parapsychological phenomena. These European investigators and students of dowsing are looking for an explanation in geophysical terms. The line of evidence which is cited as supporting this kind of hypothesis is derived from tests in which different dowsers independently work over the same area. While the review does not attempt to summarize the results in detail, the author implies that such tests have shown a degree of similarity of responses by different dowsers that cannot reasonably be attributed to chance variation. The interpretation favored is that such similarity is due to some physical influence of those locations on sensitive persons and thus accounts for the turning of the dowsing rod. Other possible interpretations which deserve consideration are that there might be some subtle visual cues which affect some of the dowsers in the same way and thus tend to elicit similar responses, or that ESP itself may be a factor which tends to produce similar responses under the circumstances holding in these tests.]

THE PROBLEMS related to the dowsing rod have been a subject of controversy since Prokop and his co-workers concluded in their book (1) that dowsing is due only to self-deception and fraud. The majority of scientists, especially the geologists, are definitely hostile to the dowser. He is to them what the quack is to the medical doctor. Actually, these days, there are a number of charlatans and dreamers who discredit the art. Whoever has an open mind on the problem and wants to hear both sides of the argument finds himself battling it out with both contenders. In 1949 (17) I very cautiously expressed my opinion on the matter, pointing out the various sources of mistakes, especially in the psychic field. I also mentioned that the occasional subjective *interpretation* of the dowsing rod reaction by dowsers without scientific training is questionable. Until now this has made it impossible to use the methods of dowsing for finding subterranean mineral deposits without having experts present at the same time. There are as yet far too many factors of uncertainty. "In dowsing reactions we do not know

whether the waves which cause the *genuine* dowser to react might have been deflected, refracted, dispersed or absorbed by one or another kind of overlying layer. Until we do know, the procedure must necessarily be of uncertain value" (**17**). Furthermore, I pointed out that as far as deposits other than water are concerned, Professors A. Wurm and K. H. Blumenberg, as well as Erster Bergrat H. Werner (**22**) have come to the conclusion as a result of observations lasting several years, that there were more than ninety percent failures. At that time I asked that an official qualifying examination be given to all dowsers. I repeated this request in 1954 (**19**) with my reasons. Such restraint seems justified after careful reading of the very critical but also very factual new examination of the problem by Mende (**11**).

After having dealt with this problem and its history for several decades, I am of the opinion (without being gifted as a dowser) that in the case of the "geopathic" gift of the genuine dowser—unfortunately there are many "dowsing-neurotics"—we are dealing with an allergy; namely, with a reflex action that is triggered by outside stimuli (from the soil), not with "occult" phantasies, and that this is a question for the natural sciences. Apart from psychogenic stimuli, whose existence cannot be denied, it must be assumed that there are geopathogenic stimuli. There are too many facts to deny such an assumption.

It is interesting to observe how the arguments for and against the dowsing rod are being repeated; for example, as far back as about 45 years ago, I proved the famous ideomotoric explanation to be insufficient. It is being used again these days as the most important part of the "proof" that everything is self-deception, if not fraud. The senior master of the Swiss geologists, Albert Heim, who has collected experiences with numerous dowsers for many years, said in 1903 (**8**): "The proof that an idea can often cause the dowsing rod to react is not proof that this is *always* the case and that there is *no other reason* for dowsing rod reactions." Heim has observed quite a number of cases "where the dowsing rod has provided first-class results, although there *was no possibility* [*for the dowser*] *to have any idea of the correct responses in advance.*"

Though most scientists are opposed to dowsing, not all are, as is demonstrated by my listing published in 1954 (**18**). I mentioned

geologists in particular because I thought their opinions on the subject would be of special interest, but I could have quoted the opinion, for example, of the winner of the Nobel Prize for Physics, J. J. Thomson (**13**, p. 3); the discoverer of local anesthetics, Carl Ludwig Schleich (**16**, p. 66); the geophysicist from Göttingen, Richard Ambronn (**20**, p. 312 ff.); or prominent psychologists like Max Dessoir (**5**), R. Sommer (**20**, p. 85 ff.), Willy Hellpach (**9**) and Alfred Lehmann. Professor of Psychology A. Lehmann of Copenhagen is the author of the authoritative work *Aberglaube und Zauberei* (*Superstition and Magic*), and in the first two editions of this work he denounces the dowsing rod as old superstition. In all later editions, however, he has revised his judgment.

The well-known psychologist Prof. Willy Hellpach (who died in 1955) can be regarded as the founder of the field of research known today as bioclimatics. Already in 1911 (in his fundamental work the sixth edition of which was published in 1950) he compared the dowsing reaction with weather sensitivity, and thus considered it a type of allergy, though this term was not known at the time (**9**). Prof. Dr. -Ing. E. Brüche (certainly no proponent of dowsing) is talking to the point when he refers to "geological hayfever." This has nothing to do with supersition! Hellpach's judgment is as follows (**9**): "No doubt there are sensitive people on whose psychic condition the content of the soil (about which they could know nothing) is not without effect." Those who react with migraine or other complaints to the presence of a weather-front are, as is known, allergic.

Now we want to go into the most important objections to the use of the dowsing rod. They are not new, by any means, and their relevance will not be disputed. A major argument concerns the reported experience that two or more dowsers, working independently in the same location and without communicating, obtain results that are said to be different from each other every time.

When such tests with dowsers are arranged, then their tasks in the field tests must be prepared in such a manner that their results can be checked and evaluated without several interpretations being possible. This was done by the Prussian Geological State Institute in December, 1920, and the results were clearly negative (**12**). It should be pointed out, however, that at that time the International

Association of Dowsers refused to participate because they found the conditions unacceptable. As a result, the dowsers who did participate may not have been qualified. The tasks given them were for the most part of a purely geological nature, such as the finding of boundaries of salt, coal, calcium deposits, etc., tasks which no doubt were new and strange to them. It is a sign of their naïveté that they thought they could tackle these tasks nevertheless. Also the tests carried out by the Swedish hydrogeologist Gunnar Ekström (which he reported in 1932) permitted the conclusion (**6**): "Different dowsers always obtain different deposits at the same location; the same dowser dowsing at the same location on various occasions obtains different deposits." By deposits are meant zones of stimulation.

Holland's Academy of Sciences arranged test series with dowsers which were reported by the Dutch physicist Professor Dr. J. Clay in 1954 (**3**). It seems strange that the whole test committee consisted of only Prof. Clay and his assistant, who carried out these tests all by himself without knowing anything about the problems. One would expect that tests that were to count would have been worked out by experts, i.e., geologists and mining officials, as was the case at the trials of the Prussian Geological State Institute. According to the Dutch professor of geology, Dr. S. W. Tromp (who is adviser on geology to the United Nations), the methods used were faulty and could be expected to lead to only negative results (**14**). There was also some misleading of the dowsers who were being tested by markers in the ground, and some tests were done inside houses, which according to my firm opinion should not have been done.

The Prussian Geological State Institute in 1922 issued a brochure about the methods of experimentation with dowsing rods and pendulums. The most important items, together with comments, have been published on pages 289 ff., in the *Handbook of Dowsing* (**20**). If this book had been read more widely, there would not have been so much nonsense written for and against dowsing. It is difficult to understand why these well-thought-out methods were not taken into consideration in setting up new tests, such as those in Holland. There is a tendency to lead dowsers up the garden path, as was done there.

In my opinion it is unfair to set suggestive traps for dowsers. Such influences affect everybody, even the most sensitive dowser, as, after all, he doesn't work with the certainty and safety of a physical instrument. In 1912 I discussed (**15**, p. 464) in great detail such a case in Britain, where a geopathically gifted gentleman reacted at the wrong places after markers had been moved secretly. Leading questions are not recognized in a court of law as a means of finding the truth.

We could also mention examples in which various dowsers in field tests have come up with *similar* results. Such an observation was made, for instance, by Walther (**21**). Geheimrat Dr. Walther was Professor of Geology at Halle University and President of the oldest German academy of sciences, the "Leopoldina" in Halle. He carried out field tests of dowsing for many years with the help of his students. Page 24 of his brochure mentions: "My students were familiar with the plotting of geological sections on graph paper. I therefore asked them to put down on strips of such graph paper the major streets of Halle's suburbs, and then skilled dowsers obtained the various stimuli along the path. Every examinee then had to walk the same path with the dowsing rod all alone, and carefully put down on the graph strips all points at which the dowsing rod had reacted. Then we got together, put the paper strips together and compared the marked points. It turned out that not everybody had found all stimulus points, some had found only part of them, but nevertheless *these points were generally in the same spots.*"

Walther found that about ten to fifteen percent were gifted "geopathically," although, of course, sensitivities varied. It is thus quite understandable that one might react to dislocated layers or other disturbance zones which are not noticed by another.

Apparently there was no intention to prove by geological means whether these results were correct, as was done by the geophysicist Dozent Dr. Volker Fritsch who reported a tectonic disturbance zone near Tetschen (**7**). On page 30 Fritsch says that the proof of such a zone of disturbance "was given by four dowsers who didn't know the location and who did their work independently of one another on four consecutive days. Their results were pretty close to each other's." This is shown by means of four sketches. Fritsch also says, on page 22: "It must be understood that the same stimu-

lus reacts in different manners on different people of different sensitivity. Therefore the reaction obtained by different persons at the same location can be quite different."

A curious but useful experience is told in the *Handbook of Dowsing* referred to earlier (**20**, pp. 275-76): "In Marienberg (Erzgebirge) some years ago a committee was created whose purpose was the finding of the lost spa (apparently lost about 100 years ago owing to displacement of the spring outlet) and make it useable again for the town of Marienberg. For this purpose four dowsers (some of them well known) who were said to be experts at this sort of work were asked to help. They went over the terrain (entirely independently of one another, one week apart) with the dowsing rod. *With surprising agreement* all four pointed out a particular spot just above the station, where drilling should commence in order to reopen the spa. The predictions of the depth to which drilling was to be done varied between 60 and 180 metres, yet there was *no sign of water* and the result therefore was negative, although some trifling indications of ore and precious metal were found."

This was a miss, but what was the cause of it? To reply to this question with answers such as coincidence or superstition is nothing but an evasion.

We are then surely entitled to the conclusion that the genuine dowser reacts to stimuli from the soil. This has also been proposed (among others) by the Swiss geohydrologist Dr. J. Kopp in many of his publications (**10**). However, the *explanation* of the reaction remains a difficult problem which can only be solved in due time by geological experts. The dowser should only be a tool, a registering instrument, in the hands of geologists. So far no positive statements can be made about the nature of the stimuli. The complex concept of "earth rays" is only a conventional expression which has become popular. Fritsch (**7**) and Mende (**11**) have discussed the existing theoretical possibilities without, however, giving any decisive and satisfying answers.

In passing, the controversial tests of the engineer Pierre Cody of Le Havre should be mentioned briefly. Cody thought he had found that the ionization of the air was a hundred times stronger at

locations where there had been cases of cancer than it was only a few meters away (**4**). Members of the Academie Française criticized this contention in 1936 (**23**). Their objections did not concern Cody's methods and the results of his measurement, but rather his assertion that there was a connection between incidence of cancer and the earth rays which he had proven by means of his measurements of the level of ionization. Whether thereby an explanation can be found for the "strips of stimuli" is another question. Mende says in this connection (**11**) that the hypothesis that the "strips of stimuli" are locations of radioactive materials and thus would explain the increased ionization of the atmosphere is not only important because of Cody's results; it might also be a plausible explanation for "nests of lightning," biological phenomena, and other observations.

As far as the physical proof of "geopathic zones" is concerned, the lectures of J. Wüst, Ph.D., M.D., "On the Actual Position of Geopathy," (**4**) and of H. Petschke, M.D., "Physical Measurements of Geopathic Zones" (**4**), as well as the recent publication of Prof. Dr. S. W. Tromp (**14**) should be mentioned.

It is not a question of measuring "earth rays." These are experimental studies of the electrical conductivity of the soil inside and outside stimulus zones as they were carried out by Tromp with two electrodes, and by Petschke according to the four-point method. The above-mentioned report by Petschke mentions twenty-eight such typical measurements which are illustrated graphically. There are a total of about five hundred along approximately twenty-five kilometres.

It is shown thereby that it is not only possible to prove that a "location" pointed out by a dowser has a greater or lesser conductivity than the surrounding area, but also to examine the course of such a zone or geological disturbance and to confirm it. It can be shown thereby that there is a real basis for the reaction of suitable dowsers. In addition to Brüche (**2**), Gärtke and Schumann had results similar to those of Wüst and Petschke.

The problem of the dowsing rod is a sore point that nobody likes to tackle. As long as the appropriate experts do not want to go into it seriously, a further clarification cannot be hoped for.

REFERENCES

1. ABELE, G., BÖCKLER, R. *et al. Wünschelrute, Erdstrahlen, und Wissenschaft* (*Dowsing-rod, Earth Rays, and Science*). Stuttgart: Otto Prokop, 1955.
2. BRÜCHE, E. *Naturwissenschaftliche Rundschau* (*Scientific Survey*). (Sept. 1954), **7**, 367-77.
3. CLAY, J. *Kon. Nederland. Akadem. van Wetenschappen* (*Royal Dutch Academy of Science*), 1954, **43**, Nos. 3 and 4.
4. CODY, PIERRE. *Erfahrungsheilkunde.* (*Experiential Medicine*), 1954, **3**, No. 12.
5. DESSOIR, MAX. *Die Woche* (*The Week*). 1906, **38**, 1637-38.
6. EKSTRÖM, G. *Slagrute och vattenador* (*Dowsing Rod and Waterveins*). Stockholm, 1932.
7. FRITSCH, VOLKER. *Das Problem Geopathogener Erscheinungen von Standpunkt der Geophysik und Ahnlicher Phänomene* (*The Problem of Geopathogenic Phenomena from the Point of View of Geophysics and Similar Occurrences*). Munich, 1955.
8. HEIM, A. *Vierteljahrsschrift der Naturforschenden Gesellschaft zu Zürich* (*Quarterly of the Zürich Science Association*), 1903, **48**, 287-306. Reprinted in *Journal für Gasbeleuchtung*, (*Journal of Gaslighting*), 1905, **48**, 1091-96.
9. HELLPACH, WILLY. *Die Geopsychischen Erscheinungen. Wetter, Klima und Landschaft in ihrem Einflusz auf das Seelenleben* (*The Geopsychic Phenomena. Weather, Climate and Location in Their Effects on the Mind*). 6th ed. 1950.
10. KOPP, J. *Oncologia* (*Oncology*), 1955, **8**, 227-33; *Verhandlungen der Schweizer Naturforschenden Gesellschaft* (*Proceedings of the Swiss Association of Science*), 1954, 147-150; *Gesundheit und Wohlfahrt* (*Health and Welfare*). 1954, No. 9, 397-401, 407-411.
11. MENDE, HERBERT G. *Die Wünschelrute und was Dahintersteckt* (*The Dowsing-rod and What Is Behind It*). Munich, 1956.
12. Preuss. Geolog. Landesanstalt (ed.). *Zur Wünschelrutenfrage* (*On the Question of Dowsing*). Berlin, 1921.
13. THOMSON, J. J. *Recollections and Reflections*, 1937, according to Rhine in *Journal of Parapsychology*, 1952, **16**, 1-10.
14. TROMP, S. W. *Experiments on the Possible Relationship Between Soil Resistivity and Dowsing Zones*, 2nd ed. Leiden: Oegstgeest, 1956.
15. VON KLINCKOWSTROEM, CARL. Etwas von Miszerfolgen der Wünschelrute (Something on the failures of dowsing), *Das Wasser*, 1912, **8**, 464.

16. ———. *Neues von der Wünschelrute* (*News About Dowsing*). 2nd ed., Berlin, 1919.

17. ———. Das Problem der Wünschelrute (The problem of dowsing), *Die Wasserwirtschaft,* 1948-49, **39**, 152-53.

18. ———. Gotachtliche Urteile über die Wünschelrute (Expert opinions about dowsing), *Die Wasserwirtschaft,* 1953-54, 99-104.

19. ———. Archiv zur Klärung der Wünschelrutenfrage (Archive for clarification of the question of dowsing), *N.F.H.*, 1954, **2**, 48-49.

20. VON KLINCKOWSTROEM-MALTZAHN. *Handbuch der Wünschelrute* (*Handbook of Dowsing*). Munich, 1931.

21. WALTHER, JOH. *Das Rätsel der Wünschelrute* (*The Riddle of the Dowsing-rod*). Leipzig, 1933.

22. WERNER, H. *Kali* (*Potash*). 1928, **22**, No. 21-24.

23. WÜST, J. *Die Heilkunst* (*The Healing Art*), 1951, **64**, 24-30.

Ainmillerstrasse 33
München 13, Germany

COMMENTS AND LETTERS

Statistical Evaluations Used in the Schmeidler and McConnell Book

To the Editors:

Dr. Thouless' review[1] of the book *ESP and Personality Patterns* by Schmeidler and McConnell is critical of certain of the statistical methods used in the book. As Dr. Greenwood and I read the manuscript with particular attention to statistical methodology, we feel a certain responsibility in this connection.

On page 221 Dr. Thouless says: "It is disappointing to find on page 46 a reversion to Dr. Schmeidler's original calculation of the significance of the difference between the mean scores of sheep and goats as evidence of the reality of the difference in scoring level of these two categories. This calculation is certainly all that is required to demonstrate the force of the evidence from these experiments for the reality of ESP, but to show that the observed difference is due to the sheep-goat difference, it is necessary to employ some method of estimating significance which treats the individual subject and not the guess as the unit."

In fact, the mathematical model used for the analysis of variance displayed in Table 9 (p. 49), on which the probability given on page 46 and cited by Thouless was based, was such that in effect the subject was treated as the unit, even though this is not entirely clear from the table itself.

Later in the same paragraph Thouless further says: ". . . on page 55 an experiment by Bevan using this method is quoted without a query. . . ." As a matter of fact, Bevan did not employ analysis of variance, but evaluated the deviation from chance expectation using the theoretical variance. In the Table of Contents of the book there appears the entry "Previously Published Data" leading to page 122, where the method of presenting the data of Bevan and others is explained and justified.

Sincerely yours,

(Signed) T. N. E. Greville

[1] *J. Parapsychol.*, **22**, p. 220-22.

To the Editors:

In the September, 1958, number of the JOURNAL, I criticized the authors of *ESP and Personality Patterns* for using on p. 46 an invalid argument for the reality of the sheep-goat difference. It has been pointed out to me by Mr. C. S. O'D. Scott that the estimate of significance here was not obtained in the way I had supposed but from an analysis of variance, which method is not open to my criticism. I wish, therefore, to be allowed to withdraw the criticism and to apologize to Dr. Schmeidler and Dr. McConnell for having made it. I do not, however, withdraw my criticism of the reference to the Bevan experiments on p. 55 which still seems to be an example of the fallacy I mistakenly supposed was present in the test of significance on p. 46. It is true that on p. 122 it is stated that the significances quoted for such experiments are "only approximate." Since one could know that they are approximate only by calculating the value properly and comparing it with the published figure, there seems no ground for supposing that they are even approximately correct. The authors' observation that "observed variances are usually within 10 per cent of theory" was presumably based on groups larger than ten, which was the number in the groups used by Bevan.

Yours sincerely,
(signed) R. H. THOULESS

BOOK REVIEWS

THE INTEGRATION OF HUMAN KNOWLEDGE. By Oliver L. Reiser. Boston: Porter Sargent, 1958. 478 pp. $8.00.

Oliver Reiser's *Integration of Human Knowledge* is a courageous book. It outlines a theme of immense appeal and calls attention to a universal task of utmost importance. As such it is a valuable contribution and deserves to be read with sympathy and care.

The literature is full of precepts for integrating knowledge. Most of them are based on the practical exigencies of our time. They deplore our moral shortcomings, the inadequacies of our educational systems, the failures of our efforts at international communication. Some writers want us to return to one or the other of our established religions, others advertise a shallow brand of scientific humanism. Reiser's approach is in this sense different from all of these and probably unique: he raises the issue of modern man's predicament to a problem of metaphysical proportions and suggests a solution as audacious and sweeping as the winds of March that break the ice on dams and rivers. The question is whether the metaphysical system proposed is sufficiently appealing and correct in its factual foundation to break man's hidebound preconceptions.

For in its very uniqueness lies a difficulty: Reiser promises salvation to those who accept his philosophic doctrine; hence conversion is a preliminary to this book's success. I do not feel secure nor informed enough to appraise the soundness of the modernistic edifice of Reiser's metaphysics in its entirety. It has features which impress me greatly; its reliance on science comforts me; its vast extrapolations beyond verifiable knowledge are hanging arches which give me dizzy spells.

But there is a small cornice, interesting to the readers of this journal, that invites my inspection and emboldens me to comment. It is Reiser's treatment of parapsychology, an attractive treatment because it endeavors to incorporate this emerging unorthodoxy within a general scheme of things.

There is perhaps nothing new in the concept of a "universal consciousness functioning through a cosmic field"; this is a feature of all panpsychisms and many pantheistic views. Nor is it very striking to note that ESP is possible because of an immersion of the human consciousness in this cosmic field. Interesting are the details with which Reiser invests this field. Some of them are quite disturbing. Almost all analogies (euphemised, perhaps for ready acceptibility, into "homomorphisms") with the fields of physical science detract from the soundness of Reiser's proposals and render them internally inconsistent. In one place we are told that the field is a "hyperdimensional medium" in which the space-time universe is suspended and which is itself outside space and time. As another it, or at least part of it, is said to act like a Heavyside layer surrounding the earth and capable of reflecting psi signals. The crucial medium is sometimes characterized as an "energy field," and yet there are profuse references to a superdispersive ether in which velocities greater than that of light are possible. No account is taken of the fact that in a superdispersive medium, energy *can not* flow with super-light velocities. Most unfortunate is the pretense at mathematical formulations which results in such atrocities as the integrals on p. 429.

Despite these infelicities, the percipient reader must acknowledge one remarkable contribution to our nascent understanding of psi phenomena. Reiser points out repeatedly that atomic physics operates with formal laws and principles, with non-dynamic mathematical relations which supervene upon the force-induced phenomena of nature. He mentions the exclusion principle as an important instance, and he might have added other kinds of symmetry requirements and postulates of invariance. He clearly sees that these relations are timeless, omnipresent and all-controlling. If these principles could be said to form a field, a logical field, indeed a logos, the partaking of human consciousness in it might at times well have the character of paranormal perceptions. To present a simple instance, the conclusion from true and present premises can break forth simultaneously in several minds, and it can be simultaneous with the happening of these conclusions or indeed precognitive. To me, this is the most interesting part of Reiser's affirmation. In the face of classical science, the instance I have chosen is trivial. But

the logos of modern science which employs abstract relations with increasing confidence and frequency, is growing in extent and importance. There may be far greater opportunities to infer, consciously or unconsciously, from the logical milieu around us than physicists and psychologists yet realize.

HENRY MARGENAU

THE WORLD OF PSYCHIC PHENOMENA. By F. S. Edsall. New York: David McKay, 1958. 244 pp.

This is a readable and rather comprehensive introduction for the general lay reader to the broad field of psychic phenomena, ranging from reported ghosts and apparitions to laboratory experimentation on extrasensory perception. The author has reviewed in a short space a large portion of the classic psychical research literature and a lesser portion of the reported experimental findings. She has made a straightforward approach and a careful attempt to set up criteria for sorting the wheat from the chaff. The thoughtful reader will welcome her discerning judgment in matters such as the Bridey Murphy case, the "contributions" of the East, mediumistic material, and the like. The evidence she does then find for a real world of parapsychical phenomena will carry that much more weight for the beginner in this field. While only twenty pages are devoted to experimental work, these go far to summarize the current research picture. Even psi-missing is discussed. The least satisfactory portion of the book, as indeed of the field itself, is that which attempts to elucidate the theories which have been proposed to explain parapsychical phenomena. But then, how could one expect to make understandable such involved ideas as Jung's synchronicity in one page?

SARA R. FEATHER

PSYCHICAL RESEARCH. By Reginald Omez (translated by Renée Haynes). New York: Hawthorn Books, 1958. 144 pp. $2.95.

Psychical Research is Volume 36 of the TWENTIETH CENTURY ENCYCLOPEDIA OF CATHOLICISM, but the seventh to be published. The reader is advised to keep in mind that Father Omez is a theo-

logian and miracle investigator for the Catholic Church: this is a primer for parapsychology aimed at a Roman Catholic audience.

Father Omez begins his book well by defining his terms (which are mostly those in current use by parapsychologists except for a few pointed out by the translator), and then outlines the history of the field with the emphasis on developments in France and Belgium. However, he gets on less firm ground when, in Chapter II under the subheading "Subject Matter," he broadly states, "The first task of parapsychology must therefore be to attempt firmly to establish the existence of its subject matter as such . . . . This is still, today, the most urgent preoccupation of all parapsychologists and psychical research societies worthy of the name." Father Omez then spends considerable time discussing the methods of evaluating spontaneous and mediumistic data. This is followed by a brief note on controlled experimentation ("qualitative method") and a longer commentary on statistical evaluation ("quantitative method"). He believes that genuine occurrences may be concealed behind statistically insignificant data and that they must not be dismissed too quickly.

In the same chapter he refers to the danger of being misled by such phenomena as psychological automatism, perceptual distortion, and fraud. He discusses with much vigor and familiarity the apparent psychical occurrences simulated by persons suffering from hysteria.

In his third chapter, "Psychical Research and Religion: Two Clearly Demarcated Fields," he lambastes parapsychologists for toying with their subject's implications for religion and philosophy. Here he goes to great length to cite the Church's view on such phenomena as the field embraces and to discuss the method of examination which the Catholic Church uses for possibly miraculous occurrences. This chapter is of interest to the theologian as well as to the student of spontaneous cases.

Chapter IV, "Conclusions or Contemporary Scientific Opinions," will certainly bring some protests from parapsychologists. Father Omez divides the field into physical and psychological phenomena and proceeds to exhibit a complete unfamiliarity with recent PK data, debunking physical effects as unproven though possible. He seems relatively familiar with telepathy but shows a lack of knowl-

edge about clairvoyance and precognition. This is by far the poorest chapter of the book.

In Chapter V, "Is Parapsychology a Science of the Future?" he offers the speculation that the psi function in man has atrophied, stating bluntly that animals show a greater psi ability. Here he also proposes that parapsychology is a small branch of classical psychology rather than a science in its own right.

This book, which began so well, finally dies with a whimper as Father Omez envisions the future role of parapsychologists as religious and scientific trouble-shooters, rooting out the frauds, charlatans, and hysterics for the churches and sciences and using parapsychological publications as exposé magazines, revealing to the public what the scientists and theologians haven't time to worry with.

W. S. Cox

PARAPSYCHOLOGICAL ABSTRACTS

SARA R. FEATHER

[This section consists of abstracts of reports published in other journals and of unpublished reports filed at the Parapsychology Laboratory at Duke University. Readers may obtain unabridged copies (Thermo-Fax) of the unpublished reports by writing to the *Journal of Parapsychology,* College Station, Durham, N. C. The price is ten cents per page.]

Ehrenburg, W. Psi-Versuche mit Uhrzeigerkarten (ESP experiments with clock cards). *Psychophysikalische Zeitschrift,* 1957, 3, 282-83.

The results of a series of impromptu clock card experiments are published with the special purpose of demonstrating the efficiency of the author's graphic-statistical method of assessment (as described on pages 220-23 and 237-40). In this experiment the usual clock cards indicating 12 hours were replaced by some in which the clock face was divided into 24 sections. The GESP technique was adopted. The nine percipients sat with the agent in the same room, facing him. Only one run through the pack of 24 clock cards was carried out. Every card in the pack had its pointer in a different position. Five percipients scored significantly above chance; the subject who scored highest had results which gave a CR of 4.6 ($P = .000{,}005$).

G. ZORAB

NELSON, R. E. Faith can move mountains. (Unpublished manuscript.)

A PK experiment is reported in which the author as sole participant attempted to win at 1000 games of "craps" more often than could be expected by chance variation alone. The total number of games won within the 1000 played was not statistically significant. However, upon learning two years after completing the experiment that declines are commonly found in PK data, the author reanalyzed his results and discovered significant decline effects. He won 270 of the first 500 games as compared with only 226 of the

last 500 games. A chi-square of the difference is statistically significant ($P = .0054$). A still more significant difference ($P = .0002$) was found upon comparing the first 250 games with the last 250 games. Analysis of 1008 throws made after completion of the experiment seemed to indicate that the dice used in the experiment were unbiased.

RYZL, MILAN. Report on experiments relating to the influence of hypnosis on ESP. (Unpublished mansucript.)

This paper presents the author's impressions concerning the relation of hypnosis to ESP on the basis of 12 years of work on that subject. From 203 unselected subjects, mostly between 15 and 30 years of age, he obtained good results under hypnosis with 24. The author attributes his inability to develop ESP in all hypnotized subjects to a failure to induce the necessary "states of consciousness." Also, he observes that successful subjects sometimes suffered an irreversible loss of ability to score well.

Due to a disappointing level of success with ESP cards, a more varied range of target objects was introduced. These included objects and scenes from daily life which did not lend themselves well to quantitative evaluation, and also some watch-face targets which were more easily assessed. The test records of two selected subjects are described by way of illustration. In 14 attempts to guess the position of both watch hands, the subjects made 10 correct hits. This is a highly improbable number of hits to be expected by chance. However, the report does not make clear the range of error allowed in scoring a hit or give the assurance needed as to the randomness of the target settings of the hands.

SERVADIO, E. Telepathy and psychoanalysis. *J. Amer. Soc. psych. Res.*, 1958, **52**, 125-33.

Telepathy and other parapsychological phenomena are interpreted by the author (as, he shows, they were by Freud) in terms of the same psychoanalytic principles that would apply to normal mental activity in general. Servadio further maintains that spontaneous telepathic occurrences between two persons, analyst and patient for instance, involve them in a pattern (a Gestalt) which

comprises them both; in analytic terms, he would call this a transference-countertransference relationship.

STRAUCH, INGE. Zur Frage der «Geistigen Heilung». (On the question of "spiritual healing.") *Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie,* 1952, **2**, 41-64. [The following summary is adapted from the original article.]

Starting from the definition of "spiritual healing" given by Thouless, the author reports the motivation, the methods, and the results of an investigation which was performed by the Medizinische Poliklinik of Freiburg University and the Freiburg Institut für Grenzgebiete der Psychologie und Psychohygiene, concerning the activity of a "spiritual healer." Six hundred and fifty patients were medically and psychologically examined and the results of the contact with the "spiritual healer" were observed up to 14 months after treatment. About 61 percent of the patients made subjective judgments indicating improvement (three-quarters of the patients were chronically ill). In contrast to these improvements there were only a few objective improvements.

Finally, the psychological structure of positively reacting patients is examined, improved methods are proposed, and the agencies of "spiritual healing" are discussed.

TENHAEFF, W. H. C. Über die Anwendung paranormaler Fähigkeiten. (On the use of paranormal abilities.) *Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie,* 1958, **2**, 10-29. [The following summary is reprinted from the original article.]

The author discusses a series of cases concerning the use of sensitives by the police or for other practical purposes. Apart from the relatively rare cases in which complete success was obtained, there are many others where the statements of the sensitives were of no practical value but proved to be of parapsychological interest. The contribution deals mainly with the depth analysis of such failures. The author follows up the motivation of the paranormal impressions in connection with the sensitive's life history. Affective factors often seem to divert the sensitives from the task in question. Laymen should not consult sensitives, for only experts can make

allowance for the many sources of errors (e.g., telepathic transference of expectations).

VAN DE CASTLE, R. L. An exploratory study of some personality correlates associated with PK performance. *J. Amer. Soc. psych. Res.*, 1958, **52**, 134-50.

Thirty-one subjects in a dice-throwing experiment achieved only chance results in 13,392 die-throws, nor did further analysis reveal any clearcut relationships between PK scoring and several personality correlates known to be related to ESP success. Some suggestion that expansive ("reaching out") subjects may score better than compressive (restrained, inhibited) subjects in PK tests was indicated by a significant difference in the scoring consistency of "completely expansive" and "completely compressive" subjects. A second suggestion in this direction comes from a marginally significant rank-order correlation which was found between the average PK run scores and the subjects' degree of expansiveness. No significant relationship appeared, however, when the PK data were analyzed with respect to the subjects' belief or nonbelief in the theoretical possibility of PK, their ratings on the Rosenzweig Picture-Frustration Test, or intelligence as measured by the Raven Progressive Matrices. A preliminary analysis of Rorschach results suggested a possible relationship between PK ability and "spontaneity."

WARCOLLIER, R. Experiences d'ESP à distance par coupe des cartes. (ESP experiments at a distance involving the cutting of the packs.) *Revue Métapsychique*, 1957, **1**, 26-31.

The experimental set-up of the series reported here was as follows: Mr. Warcollier, in Paris, acting as the GESP agent, drew an ESP card at random from a pack once a week. On that same day, Dr. and Mrs. Vasse, living at Amiens and acting as two percipients, each shuffled a pack of cards, either by hand or some mechanical device, and then cut it by hand. The card indicated by the cut was to be noted as constituting a guess. This procedure was repeated until at most 10 runs through the pack, or 250 cuts for the one target, had been made. There were 10 such weekly sessions. Dr.

Vasse's score was not significant (CR = 1); Mrs. Vasse, however, obtained a CR of 2.70.

A second series of 12 weekly sessions was started after the Vasses had visited Mr. Warcollier's house for the first time. In this series Mrs. Vasse alone acted as percipient. During one week the GESP technique was used and during the next, clairvoyance, Mr. Warcollier not looking at the card he had drawn. No above-chance results were obtained with either technique, though the GESP series was marginally significant, with a CR of − 2.13. A friend of the Vasses in Amiens, a Mr. M.D., was induced to act as a percipient in the GESP part of this series. His score, which was significant, was even more negative than Mrs. Vasse's (CR = − 2.50; P = 0.012).

The writer discusses the possible causes of the deterioration of Mrs. Vasse's score. He believes that one of these may have been the state of depression Mrs. Vasse went through during the second series.

G. Zorab

THE PARAPSYCHOLOGICAL ASSOCIATION

THE NEW COUNCIL

Elections to the Council of the Parapsychological Association were held between November 15 and December 15. The new Council consists of Dr. G. S. Schmeidler, President; Dr. Gardner Murphy, Vice-President; Mr. W. G. Roll, Secretary; Dr. R. A. McConnell, Treasurer; and three Councilmen, Miss Margaret Anderson, Dr. Karlis Osis, and Dr. J. G. Pratt.

The Membership Committee for 1958 announces that by the end of that year the Association had 51 Members and 71 Associates.

SECOND ANNUAL CONVENTION

The second annual convention of the Parapsychological Association will be held in New York City, New York, from Thursday, September 10, through Saturday, September 12, 1959. The Chairman of the Program Committee is Dr. Karlis Osis and the Chairman of the Arrangements Committee, Mr. Robert Taetzsch.

CALL FOR PAPERS FOR THE SECOND ANNUAL CONVENTION

Any Member or Associate who wishes to present a paper at the convention should write to the Chairman of the Program Committee, Dr. Karlis Osis, 80 Glenridge Avenue, Glen Ridge, New Jersey, U. S. A., to register his intention and to record the topic of his proposed paper (papers may be read *in absentia* if the member is unable to attend in person). In case a member is undecided regarding the choice of topic, two or more titles may be listed. For anyone to be eligible to present a paper his letter must reach the Chairman by May 1, 1959. The Program Committee will then send each member proposing to give a paper information regarding the preparation of abstracts. Decisions regarding the final choice of papers will be made by the Program Committee on the basis of the abstracts.

GLOSSARY

A description of the basic experimental methods, of the findings, and of the statistical procedures for evaluating ESP and PK results may be found in *Parapsychology, Frontier Science of the Mind* by J. B. Rhine and J. G. Pratt (published by Charles C. Thomas, Springfield, Illinois, U.S.A., Blackwell Scientific Publications, Ltd., Oxford, England, and The Ryerson Press, Toronto, Canada).

AGENT: The "sender" in tests for telepathy, the person whose mental states are to be apprehended by the percipient. In GESP tests, the person who looks at the stimulus object.

BT (Basic Technique): The clairvoyance technique in which each card is laid aside by the experimenter as it is called by the subject. The check-up is made at the end of the run.

CALL: The ESP symbol selected by the subject in trying to identify a target.

CHANCE: The complex of undefined causal factors irrelevant to the purpose at hand.

Chance Expectation = Mean Chance Expectation: The most likely score if only chance variation is present.

Chance Average: Mean chance expectation expressed as an expected score, generally in terms of average per run.

CHI SQUARE: A sum of quantities each of which is a deviation squared divided by an expected value. Also a sum of the squares of CR's.

CLAIRVOYANCE: Extrasensory perception of objective events as distinguished from telepathic cognition of the mental activities of another person.

CR (Critical Ratio): A measure to determine whether or not the observed deviation is significantly greater than the expected random fluctuation about the average. The CR is obtained by dividing the observed deviation by the standard deviation. (The probability of a given CR may be obtained by consulting tables of the probability integral, such as Pearsons.)

CR_d (Critical Ratio of the Difference): The observed difference between the average scores of two samples of data divided by the standard deviation of the difference.

DEVIATION: The amount an observed number of hits or an average score varies from the mean chance expectation or chance average. A deviation may be total (for a series of runs) or average (per run).

DISPLACEMENT: ESP responses to targets other than those for which the calls were intended.

Backward Displacement: ESP responses to targets preceding the intended targets. Displacement to the targets one, two, three, etc. places preceding the intended target are designated as (—1), (—2), (—3), etc.

Forward Displacement: ESP responses to targets coming later than the intended targets. Displacement to the targets one, two, three, etc. places after the intended target are designated as (+1), (+2), (+3), etc.

DT (Down Through): The clairvoyance technique in which the cards are called down through the pack before any are removed or checked.

ESP (Extrasensory Perception): Awareness of or response to an external event or influence not apprehended by sensory means.

ESP CARDS: Cards, each bearing one of the following five symbols: star, circle, square, cross and waves (three parallel wavy lines).

ESP PACK: Twenty-five ESP cards.

Closed Pack: An ESP pack composed of five each of the five symbols.

Open Pack: An ESP pack made up of the ESP symbols selected in random order.

EXPECTATION: See *Chance.*

EXTRACHANCE: Not due to chance alone.

FREE MATERIAL: The stimulus objects of experiments in which an unlimited or unspecified range of stimulus objects is employed (as contrasted with methods such as card-calling in which the subject knows that the stimulus object is one of a known range).

GESP (General Extrasensory Perception): A technique designed to test the occurrence of extrasensory perception, permitting either telepathy or clairvoyance or both to operate.

MCE (Mean Chance Expectation): see *Chance.*

P (Probability): A mathematical estimate of the expected relative frequency of a given event if chance alone were operative.

PARAPSYCHOLOGY: A division of psychology dealing with behavioral or personal effects that are demonstrably nonphysical (that is, which do not fall within the scope of known physical principles).

PERCIPIENT = SUBJECT: The person who makes the calls in an ESP test.

PK (Psychokinesis): The direct influence exerted on a physical system by a subject without any known intermediate physical energy or instrumentation.

PLACEMENT TEST: A PK technique in which the aim of the subject is to try to influence falling objects to come to rest in a designated part of the throwing surface.

PRECOGNITION: Cognition of a future event which could not be known through rational inference.

PREFERENTIAL MATCHING: A method of scoring responses to free material. A judge ranks the stimulus objects (usually pictures in sets of four) with respect to their similarity to, or association with, each response; and/or he ranks the responses with respect to their similarity to, or association with, each stimulus object.

PSI: A general term to identify personal factors or processes which are nonphysical in nature. It approximates the popular use of the word "psychic" and the technical one, "parapsychical."

PSI-MISSING: Exercise of psi ability in a way that avoids the target the subject is attempting to hit.

PSI PHENOMENA: Occurrences which result from the operation of psi. They include the phenomena of both ESP (including precognition) and PK.

PSYCHICAL RESEARCH: Older term used for parapsychology.

QD (Quarter Distribution): The distribution of hits in the record page (or in a logical subdivision thereof, such as the set or the half-set) as found in the four equal quarters formed by dividing the selected unit horizontally and vertically.

RANDOM ORDER: An order of events which displays no trends or regularities that would allow any inference regarding one event from one or more of the others in the series.

RUN: A group of trials, usually the successive calling of a deck of 25 ESP cards or symbols. In PK tests, 24 single die throws regardless of the number of dice thrown at the same time.

SCORE: The number of hits made in any given unit of trials.
Total Score: Total of scores made in a given number of runs.
Average Score: Total score divided by number of runs.

SD (Standard Deviation): Usually the theoretical root mean square of the deviations. It is obtained from the formula $\sqrt{npq}$ in which n is the number of single trials, p the probability of success per trial, and q the probability of failure.

SERIES: Several runs or experimental sessions that are grouped in accordance with a stated principle.

SESSION: A unit of an ESP or PK experiment comprising all the trials of one test occasion.

SET: A subdivision of the record page serving as a scoring unit for a consecutive group of trials, usually for the same target.

SIGNIFICANCE: A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. The criterion commonly used in parapsychology today is a probability value of .01 or less, or a deviation in either direction such that the CR is 2.58 or greater.

SINGLES TEST: A PK technique in which the aim of the subject is to try to influence dice to fall with a specified face up.

SPONTANEOUS PSI EXPERIENCE: Natural, unplanned occurrence of an event or experience that seems to involve parapsychical ability.

STIMULUS OBJECT: The ESP card or drawing or other object, some identifying characteristic of which is to be apprehended by the percipient.

STM (Screened Touch Matching): An ESP card test in which the subject indicates in each trial (by pointing to one of five key positions) what he thinks the top card is in the inverted pack held by the experimenter behind a screen. The card is then laid opposite that position.

SUBJECT: The person who is experimented upon. In ESP tests, most commonly the percipient (though also the agent in GESP and telepathy). In PK tests, any individual whose task it is to influence the objects thrown.

TARGET: In ESP (including precognition) the objective or mental events the subject is attempting to respond to. In PK tests, the faces of the die which the subject attempts to bring up in the act of throwing. In placement tests, the designated part of the throwing surface.

Target Card: The card which the percipient is attempting to identify or otherwise indicate a knowledge of.

Target Pack: The pack of cards the order of which the subject is attempting to identify.

Target Face: The face on the die which the subject tries to turn up by PK.

TELEPATHY: Extrasensory perception of the mental activities of another person (as distinguished from clairvoyant cognition of objective events).

TRIAL: In ESP tests, a single attempt to identify a stimulus object. In PK tests, a single throw of the dice or other objects thrown.